

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2005, 08:28:14 ; Search time 80.7044 Seconds
(without alignments)
2741.199 Million cell updates/sec

Title: US-10-017-479A-2

Perfect score: 2994

Sequence: 1 MEIEIGEQPPVKCNFFA.....NSFPWAQIYAAALGNKTH 572

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : A_Geneseq_16Dec04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2994	100.0	572	4	ABE60315
2	2994	100.0	572	4	ABE66903
3	2994	100.0	572	5	ABE79611
4	2994	100.0	572	8	ADP64789
5	2986	99.7	572	6	ABR40099
6	1327.5	44.3	908	4	ABE61068
7	1042.5	34.8	587	7	AAE38766
8	1039	34.7	581	8	ADP64799
9	1026	34.3	592	6	ABE82950
10	1026	34.3	592	6	ABR40100
11	1026	34.3	592	6	ABR57025
12	1026	34.3	592	7	ADP63755
13	1026	34.3	592	7	ADP63755
14	1010.5	33.8	619	8	ADH42443
15	1005	33.6	576	8	ADH42443
16	1004	33.5	568	6	ABR57023
17	1004	33.5	568	7	AAE38764
18	1004	33.5	568	7	ADK51052
19	1004	33.5	568	8	ADH42441
20	1004	33.5	568	8	ADP64793
21	1002	33.5	568	5	ABU65062
22	1002	33.5	568	6	ABR40097
23	1002	33.5	568	7	ADK51050
24	1002	33.5	568	8	ADH42445
25	1002	33.5	568	8	ADN61775

Ade07994	Novel pro
Aau79946	Human tra
Abg75835	Transport
ADG16993	Human tra
ABU65064	Human NOV
ABU65063	Human NOV
ADK51048	Human NOV
ADH42447	Novel hum
ADN61777	Human nov
ADN61779	Human nov
ADP64791	Rat INDY,
ABG75826	Transport
ABG82952	Human SLC
ABR40101	Human ren
ABG75837	Transport
Aae21181	Human TRI
ABR57024	Mouse TCH
ADP64797	Mouse Na+
ADR09550	Human pro
Aab36167	Novel hum
Aab36161	Novel hum
Aab36164	Novel hum
Aab36168	Novel hum
Aab23625	Human sec
Aab36158	Novel hum
Aab36162	Novel hum
Aab42213	Human ORF
ABG7450	Novel hum
Aab36165	Novel hum
Aab36159	Novel hum
ABM84184	Human dia
Aae22910	Human tra
Aao21807	Lung-spec
ABR82949	Human SLC
Adi21045	Novel hum
ADG78130	Human SLC
ADH22543	Human tra
ABM84303	Human dia
ADP29692	Human sec
ABU65065	Human NOV
ADK51054	Human NOV
ADH42449	Novel hum
ADN61781	Human nov
ADN05177	Human pro
ABU65066	Human NOV
ADK51056	Human NOV
ADH42451	Novel hum
ADN61783	Human nov
ADP64795	Nematode
Aab36169	Novel hum
Aab36163	Novel hum
Aab36166	Novel hum
Aab36160	Novel hum
Aao31005	Human tra
Aag13103	Arabidops
Aag13102	Arabidops
Aag51055	Arabidops
Aag51054	Arabidops
Aag90000	C Glutami
ABG76767	Corynebac
AAG13104	Arabidops
Aag51056	Arabidops
ADH48682	Rice prot
Adc08241	Rice prot
ABM71452	Staphyloc
Aau91115	Human-sec
ABG65198	Human alb
ADL78465	Albumin f
Adc09041	Novel pro
Aau91090	Human sec
ABG65200	Human alb
ADL78467	Albumin f

99	542	18.1	502	6	ADA48226	Ada48226 Rice prot
100	478	16.0	304	5	ABB89646	Abb89646 Human pol
ALIGNMENTS						
RESULT 1						
ABB60315						
ID	ABB60315 standard; protein; 572 AA.					
XX	AC	ABB60315;				
XX	DT	26-MAR-2002 (first entry)				
XX	DE	Drosophila melanogaster polypeptide SEQ ID NO 7737.				
XX	KW	Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical.				
XX	OS	Drosophila melanogaster.				
XX	PN	WO200171042-A2.				
XX	PD	27-SEP-2001.				
XX	PF	23-MAR-2001; 2001WO-US009231.				
XX	PR	23-MAR-2000; 2000US-0191637P.				
XX	PR	11-JUL-2000; 2000US-00614150.				
XX	PA	(PEKE) PE CORP NY.				
XX	PI	Venter JC, Adams M, Li PWD, Myers EW;				
XX	PI	WPI; 2001-656860/75.				
XX	DR	N-PSDB; ABL04418.				
XX	PT	New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signaling and cell-cell interactions.				
XX	PS	Disclosure; SEQ ID NO 7737; 21pp + Sequence Listing; English.				
XX	CC	The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-ABB72072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences				
XX	Sequence	572 AA;				
Query Match						
Best Local Similarity		100.0%;		Score 2994;	DB 4;	Length 572;
Matches 572;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
QY	1	MEIEIGEQPPVKCSNPFANHWKGLVFLVPLLCPLVPLLNEGAEFRCMYLLVMAIFW	60			
DB	1	MEIEIGEQPPVKCSNPFANHWKGLVFLVPLLCPLVPLLNEGAEFRCMYLLVMAIFW	60			
QY	61	VTEALPLYVTSMPPIVAPFIMGIMSSDQTCRLYFKDTLVMFMGGIMVALAVEYCNLHKRL	120			
DB	61	VTEALPLYVTSMPPIVAPFIMGIMSSDQTCRLYFKDTLVMFMGGIMVALAVEYCNLHKRL	120			
QY	121	ALRVIOIVGCSRRRLHFGIMVTMFLSWISNAACTAMMCPIIOAVLBELOAQGVCKINH	180			
DB	121	ALRVIOIVGCSRRRLHFGIMVTMFLSWISNAACTAMMCPIIOAVLBELOAQGVCKINH	180			

QY	181	EPQYQIVGNGKKNNEDEPPYPTKITLCYYLGIAYASSLGGCGTIIGTATNLTFFKGIYEAR	240
DB	181	EPQYQIVGNGKKNNEDEPPYPTKITLCYYLGIAYASSLGGCGTIIGTATNLTFFKGIYEAR	240
QY	241	FKNSTEQMDPFTFMFYSVPMSLVYLLTFVFLQHFHMGFWLWRPKSKEAQVQREGADVA	300
DB	241	FKNSTEQMDPFTFMFYSVPMSLVYLLTFVFLQHFHMGFWLWRPKSKEAQVQREGADVA	300
QY	301	KKVIDORYKDLGPMSEIHQIMLFIEMVVMYFTRKPGIFLGWADLLNSKOIRNSMPTIF	360
DB	301	KKVIDORYKDLGPMSEIHQIMLFIEMVVMYFTRKPGIFLGWADLLNSKOIRNSMPTIF	360
QY	361	VVMCFMLPANYAFRLRYCTRRGGPVPTGTPSLITWKFIOQKVPWGLVFLGGGFPALAE	420
DB	361	VVMCFMLPANYAFRLRYCTRRGGPVPTGTPSLITWKFIOQKVPWGLVFLGGGFPALAE	420
QY	421	SKQSGMAKLIGNALIGLVKLVNSVLLVVLVAVFLTAFSSNVAIANIIPVLAEMSLAI	480
DB	421	SKQSGMAKLIGNALIGLVKLVNSVLLVVLVAVFLTAFSSNVAIANIIPVLAEMSLAI	480
QY	481	EIHPLYLILPAGLACSMAPHLPVSTPPNALVAGYANIRTKDMAIAGIGPTIIITLFPV	540
DB	481	EIHPLYLILPAGLACSMAPHLPVSTPPNALVAGYANIRTKDMAIAGIGPTIIITLFPV	540
QY	541	CQTWGLVVPYNLNSPPEWAQIYAAAALGNKTH	572
DB	541	CQTWGLVVPYNLNSPPEWAQIYAAAALGNKTH	572
RESULT 2			
ABB66903			
ID	ABB66903	standard; protein; 572 AA.	
XX	AC	ABB66903;	
XX	DT	26-MAR-2002 (first entry)	
XX	DE	Drosophila melanogaster polypeptide SEQ ID NO 27501.	
XX	KW	Drosophila; developmental biology; cell signalling; insecticide;	
XX	KW	pharmaceutical.	
XX	OS	Drosophila melanogaster.	
XX	PN	WO200171042-A2.	
XX	PD	27-SEP-2001.	
XX	PF	23-MAR-2001; 2001WO-US009231.	
XX	PR	23-MAR-2000; 2000US-0191637P.	
XX	PR	11-JUL-2000; 2000US-00614150.	
XX	PA	(PEKE) PE CORP NY.	
XX	PI	Venter JC, Adams M, Li PWD, Myers EW;	
XX	DR	WPI; 2001-656860/75.	
XX	DR	N-PSDB; ABL11006.	
XX	PT	New isolated nucleic acid detection reagent for detecting 1000 or more	
XX	PT	genes from Drosophila and for elucidating cell signalling and cell-cell	
XX	PT	interactions.	
XX	PS	Disclosure; SEQ ID NO 27501; 21pp + Sequence Listing; English.	
XX	CC	The invention relates to an isolated nucleic acid detection reagent	
XX	CC	capable of detecting 1000 or more genes from Drosophila. The invention is	
XX	CC	useful in developmental biology and in elucidating cell signalling and	
XX	CC	cell-cell interactions in higher eukaryotes for the development of	
XX	CC	insecticides, therapeutics and pharmaceutical drugs. The invention	
XX	CC	discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA	
XX	CC	sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-	
XX	CC	ABB72072). The sequence data for this patent did not form part of the	
XX	CC	printed specification, but was obtained in electronic format directly	
XX	CC	from WIPO at ftp.wipo.int/pub/published_pct_sequences	

CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 572 AA;

Query Match 100.0%; Score 2994; DB 4; Length 572;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 572; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEIEIGEQQPPVKCSNFFANHWKGLVFLVPLLCPLVMLNEGAEFCWYLLLVMAIFW 60
DB 1 MEIEIGEQQPPVKCSNFFANHWKGLVFLVPLLCPLVMLNEGAEFCWYLLLVMAIFW 60
QY 61 VTEALPLYVTSMIPVAFPPIMGSSDQTCRLYFKDTLVFMFGGIMVALAVEYCNLHKRL 120
DB 61 VTEALPLYVTSMIPVAFPPIMGSSDQTCRLYFKDTLVFMFGGIMVALAVEYCNLHKRL 120
QY 121 ALRVQIVGCSRRRLHFGIMVTMFLSMWISNAACTAMWCPIIOAVLELOAQGVCKINH 180
DB 121 ALRVQIVGCSRRRLHFGIMVTMFLSMWISNAACTAMWCPIIOAVLELOAQGVCKINH 180
QY 181 EPOYQIVGCKNKNNEDEPPYPTKITLCYVLGIAYASSLGCGTIIGTATNLTFFKGIYER 240
DB 181 EPOYQIVGCKNKNNEDEPPYPTKITLCYVLGIAYASSLGCGTIIGTATNLTFFKGIYER 240
QY 241 FKNSTEQMDPFTFMFYSVPSMLVYLLTFVFLQWFMGLMRPKSKEAQEVQREGADVA 300
DB 241 FKNSTEQMDPFTFMFYSVPSMLVYLLTFVFLQWFMGLMRPKSKEAQEVQREGADVA 300
QY 301 KKVLDQRYKDLGPMSTHIEIQVMILFIWVYVYFTRKPGIFLGWADLLNSKDIRNSMPTIF 360
DB 301 KKVLDQRYKDLGPMSTHIEIQVMILFIWVYVYFTRKPGIFLGWADLLNSKDIRNSMPTIF 360
QY 361 VVWCMFLPANYAFRLRYCTRRGGPVPTGPTPSLITWKFIQTKVPWGLVFLGGGFALAE 420
DB 361 VVWCMFLPANYAFRLRYCTRRGGPVPTGPTPSLITWKFIQTKVPWGLVFLGGGFALAE 420
QY 421 SKQSGMAKLIGNALIGLKVLPNSVLLVILVAVFLTAFSSNVAIANIIPVLAEMSLAI 480
DB 421 SKQSGMAKLIGNALIGLKVLPNSVLLVILVAVFLTAFSSNVAIANIIPVLAEMSLAI 480
QY 481 EIHPLYLILPAGLACSMAPHLVSTPPNALVAGYANIRTKDMAIAGIGTIIITLTFV 540
DB 481 EIHPLYLILPAGLACSMAPHLVSTPPNALVAGYANIRTKDMAIAGIGTIIITLTFV 540
QY 541 CQTWGLVVPNLNSPPEWAQIYAAAALGNKTH 572
DB 541 CQTWGLVVPNLNSPPEWAQIYAAAALGNKTH 572

RESULT 3
ID ABB79611
XX ABB79611 standard; protein; 572 AA.
AC ABB79611;
XX 21-OCT-2002 (first entry)
DE Drosophila INDY protein dicarboxylate transporter.
XX Indy; dicarboxylate transporter; life span; longevity; obesity;
KW anorectic; caloric restriction; transgenic animal; gene therapy.
OS Drosophila melanogaster.
XX WO200259310-A2.
PN 01-AUG-2002.
XX 12-DEC-2001; 2001WO-US048130.
PF 12-DEC-2000; 2000US-0255013P.
PR

XX PA (UYCO-) UNIV CONNECTICUT.
XX PI Rogina B, Reenan RA, Helfand SL;
XX DR WPI: 2002-599787/64.
XX DR N-PSDB; A8N84439.
XX PT Indy polynucleotide, useful for diagnosing or treating body weight
XX disorders, e.g. obesity, metabolic maintenance disorders, or the symptoms
XX of aging to extend the life span of an organism.
PS Claim 14; Page 81-83; 83pp; English.

CC The present sequence is the protein sequence of the Drosophila
CC melanogaster INDY protein, which is encoded by the Indy gene (see
CC ABN84439) involved in increased life span. The protein has similarity to
CC dicarboxylate transporters such as those from human and rat, and
CC represents a new class of dicarboxylate transporters that are not
CC inhibited by phloretin. Identification of the Indy gene resulted from the
CC observation that particular mutations in the gene caused an increase in
CC the life span of the fly carrying the mutation. As a result of this
CC finding, it is now possible to identify and/or isolate Drosophila lines
CC with longer life spans, as well as to identify agents that contribute to
CC longer life span. It is also possible to isolate genes involved in, and
CC which have an effect on longevity, as well as proteins encoded by these
CC genes. The invention provides Indy polynucleotides, proteins, anti-INDY
CC antibodies, antagonists that inhibit Indy activity or expression and
CC agonists that increase Indy activity or expression, and their use in the
CC diagnosis or treatment of body weight disorders, such as obesity and
CC metabolic maintenance disorders, or longevity in humans and animals.
CC Antagonists include at least a portion of the Indy gene sequence, an
CC antisense oligonucleotide, a ribozyme, a triple helix-forming molecule, a
CC double-stranded interfering RNA, an anti-Indy antibody, or a mixture of
CC these. Methods of calorically restricting an organism and of extending
CC the lifespan of an organism by administering the antagonist are claimed

XX SQ Sequence 572 AA;

Query Match 100.0%; Score 2994; DB 5; Length 572;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 572; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEIEIGEQQPPVKCSNFFANHWKGLVFLVPLLCPLVMLNEGAEFCWYLLLVMAIFW 60
DB 1 MEIEIGEQQPPVKCSNFFANHWKGLVFLVPLLCPLVMLNEGAEFCWYLLLVMAIFW 60
QY 61 VTEALPLYVTSMIPVAFPPIMGSSDQTCRLYFKDTLVFMFGGIMVALAVEYCNLHKRL 120
DB 61 VTEALPLYVTSMIPVAFPPIMGSSDQTCRLYFKDTLVFMFGGIMVALAVEYCNLHKRL 120
QY 121 ALRVQIVGCSRRRLHFGIMVTMFLSMWISNAACTAMWCPIIOAVLELOAQGVCKINH 180
DB 121 ALRVQIVGCSRRRLHFGIMVTMFLSMWISNAACTAMWCPIIOAVLELOAQGVCKINH 180
QY 181 EPOYQIVGCKNKNNEDEPPYPTKITLCYVLGIAYASSLGCGTIIGTATNLTFFKGIYER 240
DB 181 EPOYQIVGCKNKNNEDEPPYPTKITLCYVLGIAYASSLGCGTIIGTATNLTFFKGIYER 240
QY 241 FKNSTEQMDPFTFMFYSVPSMLVYLLTFVFLQWFMGLMRPKSKEAQEVQREGADVA 300
DB 241 FKNSTEQMDPFTFMFYSVPSMLVYLLTFVFLQWFMGLMRPKSKEAQEVQREGADVA 300
QY 301 KKVLDQRYKDLGPMSTHIEIQVMILFIWVYVYFTRKPGIFLGWADLLNSKDIRNSMPTIF 360
DB 301 KKVLDQRYKDLGPMSTHIEIQVMILFIWVYVYFTRKPGIFLGWADLLNSKDIRNSMPTIF 360
QY 361 VVWCMFLPANYAFRLRYCTRRGGPVPTGPTPSLITWKFIQTKVPWGLVFLGGGFALAE 420
DB 361 VVWCMFLPANYAFRLRYCTRRGGPVPTGPTPSLITWKFIQTKVPWGLVFLGGGFALAE 420
QY 421 SKQSGMAKLIGNALIGLKVLPNSVLLVILVAVFLTAFSSNVAIANIIPVLAEMSLAI 480

Db 421 SKOSGMAKLIQNALIGLKVLPSVLLVWLVAVFLTAFSSNVAIANIIPVLAEMSLAI 480
QY 481 EIHPYLIILPAGLACSMAPHLPVSTPPNALVAGYANIRTKDMAIAGIGPTIITITLFFVF 540
Db 481 EIHPYLIILPAGLACSMAPHLPVSTPPNALVAGYANIRTKDMAIAGIGPTIITITLFFVF 540
QY 541 CQTWGLVVPNLNSFPPEWAQIYAAAALGNKTH 572
Db 541 CQTWGLVVPNLNSFPPEWAQIYAAAALGNKTH 572

RESULT 4
ADP64789
ID ADP64789 standard; protein; 572 AA.
XX AC
XX ADP64789;
XX DT 26-AUG-2004 (first entry)
XX DE Drosophila INDY, an Na+-coupled citrate transporter protein.
XX KW sodium-coupled citrate transporter; transmembrane citrate transporter;
KW lifespan; weight reduction; weight gain prevention; blood cholesterol;
KW triglyceride; low density lipopolysaccharide; glucose; obesity;
KW hyperlipidemia; hypercholesterolemia; INDY protein.
XX OS Drosophila melanogaster.
XX PN WO2004048925-A2.
XX PD 10-JUN-2004.
XX PF 20-NOV-2003; 2003WO-US037054.
XX PR 22-NOV-2002; 2002US-0428469P.
XX PR 01-APR-2003; 2003US-0459441P.
XX PA (MEDI-) MEDICAL COLLEGE GEORGIA RES INST.
PA (GANA/) GANAPATHY V.
PA (INOU/) INOUE K.
PA (FEIY/) FEI Y.
XX PI Ganapathy V, Inoue K, Fei Y;
XX WIPI; 2004-460797/43.
XX DR N-PSDB; ADP64788.
XX PT New isolated polynucleotide encoding a Na+-coupled citrate transporter
PT (NACT) polypeptide, useful as a drug target for the treatment of obesity,
PT hyperlipidemia, and hypercholesterolemia.
XX PS Claim 37; SEQ ID NO 2; 186pp; English.
XX CC The invention relates to novel Na+-coupled citrate transporter proteins
CC and their encoding genes. Inhibitors of transmembrane citrate
CC transporters are useful for extending the lifespan, reducing weight, LDL
CC preventing weight gain or lowering blood cholesterol, triglyceride, LDL
CC or glucose levels in a subject. The NACT polypeptide is useful as a drug
CC target for the treatment of obesity, hyperlipidemia, and
CC hypercholesterolemia. This sequence corresponds to the Drosophila INDY
CC protein.
XX SQ Sequence 572 AA;

Query Match 100.0%; Score 2994; DB 8; Length 572;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 572; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEIETGEQPPVKCSNFFANHWKGLVFLVPLLCPLVMLLNEGAEFRCMYLLVMAIFW 60
Db 1 MEIETGEQPPVKCSNFFANHWKGLVFLVPLLCPLVMLLNEGAEFRCMYLLVMAIFW 60
QY 61 VTEALPLYVTSMTPIVAFPPINGIMSSDOTCRLYFKDTLVFMFGGIMVALAVEYCNLHKRL 120

Db 61 VTEALPLYVTSMTPIVAFPPINGIMSSDOTCRLYFKDTLVFMFGGIMVALAVEYCNLHKRL 120
QY 121 ALRVIQIVGCSPPRLHFGILINVTWFLSMWISNAACATAMCPIIOAVLELOAQGVCKINH 180
Db 121 ALRVIQIVGCSPPRLHFGILINVTWFLSMWISNAACATAMCPIIOAVLELOAQGVCKINH 180
QY 181 EPOYQIVGNGKKNNEDEPPYPTKITLCYVLGIAVASSLGCGGTIIGTATNLTFFKIYEAR 240
Db 181 EPOYQIVGNGKKNNEDEPPYPTKITLCYVLGIAVASSLGCGGTIIGTATNLTFFKIYEAR 240
QY 241 FKNSTEQMDPFTFMEYSVPSMLVYLLTFVFLQWHFMGLWRPKSKEAOEVOGRGADVA 300
Db 241 FKNSTEQMDPFTFMEYSVPSMLVYLLTFVFLQWHFMGLWRPKSKEAOEVOGRGADVA 300
QY 301 KKVIDORYKDLGPMSEIHEIQWMLFIEMVVMVFTKPGIFLGWADLLNSKDIRNSMPTIF 360
Db 301 KKVIDORYKDLGPMSEIHEIQWMLFIEMVVMVFTKPGIFLGWADLLNSKDIRNSMPTIF 360
QY 361 VVVMCFMPLPANYAFRLRYCTRRGGPVPTGPTPSLITWKFQTKVPMGLVFLGGGFALAEG 420
Db 361 VVVMCFMPLPANYAFRLRYCTRRGGPVPTGPTPSLITWKFQTKVPMGLVFLGGGFALAEG 420
QY 421 SKOSGMAKLIQNALIGLKVLPSVLLVWLVAVFLTAFSSNVAIANIIPVLAEMSLAI 480
Db 421 SKOSGMAKLIQNALIGLKVLPSVLLVWLVAVFLTAFSSNVAIANIIPVLAEMSLAI 480
QY 481 EIHPYLIILPAGLACSMAPHLPVSTPPNALVAGYANIRTKDMAIAGIGPTIITITLFFVF 540
Db 481 EIHPYLIILPAGLACSMAPHLPVSTPPNALVAGYANIRTKDMAIAGIGPTIITITLFFVF 540
QY 541 CQTWGLVVPNLNSFPPEWAQIYAAAALGNKTH 572
Db 541 CQTWGLVVPNLNSFPPEWAQIYAAAALGNKTH 572

RESULT 5
ABR40099
ID ABR40099 standard; protein; 572 AA.
XX AC ABR40099;
XX DT 24-JUL-2003 (first entry)
XX DE dIndyaa sequence.
XX KW Anti-diabetic; anorectic; sodium dicarboxylate cotransporter;
KW sodium tricarboxylate cotransporter; HepNaDc; hepatocyte; HepG2;
KW diabetes; obesity; lipid metabolism; aging; dIndyaa.
XX OS Unidentified.
XX PN WO2003029465-A1.
XX PD 10-APR-2003.
XX PF 27-SEP-2002; 2002WO-JP010038.
XX PR 28-SEP-2001; 2001JP-00299433.
XX PR 28-AUG-2002; 2002JP-00249016.
XX PA (SAKA) OTSUKA PHARM CO LTD.
XX PI Kanemoto N, Omori Y, Sugano S, Obuchi Y;
XX WIPI; 2003-354728/33.
XX PT Sodium di- or tricarboxylate cotransporter gene (HepNaDC) for treatment
XX and prevention of diabetes, obesity, for improving lipid metabolism and
XX aging.
XX PS Example 1; Fig 1; 75pp; Japanese.
XX

CC The present invention relates to human sodium di- or tricarboxylate
CC cotransporter protein (HepNadC; AB400977). HepNadC gene is expressed in
CC human hepatocyte cancer-origin cells HepG2. Compounds which control the
CC expression of the HepNadC gene and activity of the HepNadC polypeptide
CC may be useful for treating and preventing diabetes, obesity, for
CC improving lipid metabolism and aging. The present sequence is the dindvaa
CC protein, which was used in a sequence alignment with the HepNadC sequence
XX
SQ Sequence 572 AA;
Query Match 99.7%; Score 2986; DB 6; Length 572;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 571; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MEIEIGEQQPPVKCSNFFANHWKGLVFLVLLCLPVMLNEGAFRCMYLLLVNAIFW 60
DB 1 MEIEIGEQQPPVKCSNFFANHWKGLVFLVLLCLPVMLNEGAFRCMYLLLVNAIFW 60
QY 61 VTEALPLYVTSMPIVAFPIGMSDQTCRLYFKDTLVFMFGGIMVALAVYCNLHKRL 120
DB 61 VTEALPLYVTSMPIVAFPIGMSDQTCRLYFKDTLVFMFGGIMVALAVYCNLHKRL 120
QY 121 ALRVQIVGCSRRRLHFGIMVTFMFLSMWISNAACCTAMMCPPIQAVLELQAGVCKINH 180
DB 121 ALRVQIVGCSRRRLHFGIMVTFMFLSMWISNAACCTAMMCPPIQAVLELQAGVCKINH 180
QY 181 EPQYQIVGKNKNNDEPPYPTKITLCYYLGIAYASSLGGCGTIIGTATNLTFKGIYEAR 240
DB 181 EPQYQIVGKNKNNDEPPYPTKITLCYYLGIAYASSLGGCGTIIGTATNLTFKGIYEAR 240
QY 241 FKNSTEQMDPPTFMFYSVPSMLVYLLTFVFLQWFMGLWRPKSKEAQVQRGREGADVA 300
DB 241 FKNSTEQMDPPTFMFYSVPSMLVYLLTFVFLQWFMGLWRPKSKEAQVQRGREGADVA 300
QY 301 KKVVDQRYKDLGPMSTHEIQVMILFIEMVVMYFTRKPGIFLGWADLLNSKDIRNSMPTIF 360
DB 301 KKVVDQRYKDLGPMSTHEIQVMILFIEMVVMYFTRKPGIFLGWADLLNSKDIRNSMPTIF 360
QY 361 VVMCFMFLPANYAFRLYCTRGGPVPTGTPSLITWKFIQTKVPWGLVFLGGGFALAE 420
DB 361 VVMCFMFLPANYAFRLYCTRGGPVPTGTPSLITWKFIQTKVPWGLVFLGGGFALAE 420
QY 421 SKQSGMAKLIGNALIGLKVLPNSVLLVILVAVFLTAFSSNVAIANIIPVLAEMSLAI 480
DB 421 SKQSGMAKLIGNALIGLKVLPNSVLLVILVAVFLTAFSSNVAIANIIPVLAEMSLAI 480
QY 481 EIHPLYLILPAGLACSMAPHLVPSTPPNALVAGYANIRTKDMAIAGIGTITITILFVF 540
DB 481 EIHPLYLILPAGLACSMAPHLVPSTPPNALVAGYANIRTKDMAIAGIGTITITILFVF 540
QY 541 CQTWGLVVPNLNSPPEWAQIYAAAALGNKTH 572
DB 541 CQTWGLVVPNLNSPPEWAQIYAAAALGNKTH 572
RESULT 6
ABB61068
ID ABB61068 standard; protein; 908 AA.
XX AC ABB61068;
XX
XX 26-MAR-2002 (first entry)
XX Drosophila melanogaster polypeptide SEQ ID NO 9996.
DE Drosophila melanogaster.
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX Drosophila melanogaster.
OS
XX WO200171042-A2.
PN
XX 27-SEP-2001.
PD

XX 23-MAR-2001; 2001WO-US009231.
PF
XX 23-MAR-2000; 2000US-0191637P.
PR
XX 11-JUL-2000; 2000US-00614150.
XX (PEKE) PE CORP NY.
PA
XX Venter JC, Adams M, Li PWD, Myers EW;
PI
XX MPI; 2001-656860/75.
DR N-PSDB; ABL05171.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
PS Disclosure; SEQ ID NO 9996; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB116176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 908 AA;
Query Match 44.3%; Score 1327.5; DB 4; Length 908;
Best Local Similarity 46.8%; Pred. No. 5.6e-136;
Matches 253; Conservative 99; Mismatches 135; Indels 54; Gaps 7;
QY 22 HWKGLVFLVLLCLPVMLNEGAFRCMYLLLVNAIFWVTEALPLYVTSMPIVAFPIW 81
DB 418 HWKGLVFLVLLCLPVMLNEGAFRCMYLLLVNAIFWVTEALPLYVTSMPIVAFPIW 81
QY 82 GIM--SSDQTCRLYFKDTLVFMFGGIMVALAVEYCNLHKRLALRVQIVGCSRRRLHGL 139
DB 451 DILESTSEKVSFYSFDVTVMFVIGGLLITALEYCNLHKRLALRVQIVGCSRRRLHGL 510
QY 140 IMVTFMFLSMWISNAACCTAMMCPPIQAVLELQAGVCKINHHEPOVQIVGKNKNNDEPP 199
DB 511 VMTCTSLWISNAACCTAMMCPPIQAVLELQAGVCKINHHEPOVQIVGKNKNNDEPP 564
QY 200 YPTKITLCYYLGIAYASSLGGCGTIIGTATNLTFKGIYEARFKNSTEQMDPPTFMFYSVP 259
DB 565 HPSTISMAFYPGIAYSSSISGCGTIGTATNLTFKGIYEARFKNSTEQMDPPTFMFYSVP 624
QY 260 -SMLVYLLTFVFLQWFMGLWRPKSKEAQVQRGREGADVAKKVIDORYKDLGPMSTHE 318
DB 625 FVLVLIPLTYFSLQVTHMGLFRPNSKIGQVKKGAESQDVVKDVIKORKAELGPMSCHE 684
QY 319 IQVMILFIEMVVMYFTRKPGIFLGWADLLNSKDIRNSMPTIFVVMCFMFLPANYAFRLY 378
DB 585 IQVGLLVILVAVFLTAFSSNVAIANIIPVLAEMSLAIEHPLYLILPAGLACSM 744
QY 379 TRRGVPVPTGTPSLITWKFIQTKVPWGLVFLGGGFALAEQSGKSNKALIGNALIGLK 438
DB 745 CGK-APFPQTLDACLSW-----GFALAEQSGRVSGWAKMGESLAFAG 786
QY 439 VLPNSVLLVILVAVFLTAFSSNVAIANIIPVLAEMSLAIEHPLYLILPAGLACSM 498
DB 787 EMHSLVISMCIISLFTAFASNAICNLIPIFSEMAIAIEVHPMKLTTPPALACSLA 846
QY 499 FHLVSTPPNALVAGYANIRTKDMAIAGIGTITITILFVFCQTKWGLVVPNLNSPPEW 558
DB 847 FHLVSTPPNALVAGYANIRTKDMAIAGIGTITITILFVFCQTKWGLVVPNLNSPPEW 906
QY 559 A 559

Db	907 A 907	
RESULT 7		
AAE38766		
ID	AAE38766 standard; protein; 587 AA.	
XX		
AC	AAE38766;	
XX		
DT	18-DEC-2003 (first entry)	
XX		
DE	Rat 69624 protein sodium sulphate symporter domain.	
XX		
XX	Rat; 69624; transporter protein; neurological disorder; therapy;	
KW	atherosclerosis; cardiac hypertrophy; ischaemia reperfusion injury;	
KW	metastatic disorder; haematopoietic neoplastic disorder; leukaemia;	
KW	arthritis; multiple sclerosis; encephalomyelitis; myasthenia gravis;	
KW	carcinoma; cell proliferation; autoimmune disorder; diabetes mellitus;	
KW	renal disorder; colon; hepatic disorder; hypocalcaemia; calcium stone;	
KW	mental retardation; Canavan disease; differentiative disorder; sarcoma;	
KW	systemic lupus erythematosus; cardiovascular disorder; arteriosclerosis;	
KW	atrial fibrillation; forensic identification; pain.	
XX		
OS	Rattus norvegicus.	
XX		
PN	US2002193582-A1.	
XX		
PD	19-DEC-2002.	
XX		
PF	17-JUN-2002; 2002US-00173519.	
XX		
PR	18-JUN-2001; 2001US-0298970P.	
XX		
PA	(MILL-) MILLENNIUM PHARM INC.	
XX		
PI	Curtis RAJ;	
XX		
DR	WPI; 2003-644620/61.	
XX		
PT	Novel 69624 polypeptide, a human transporter family member, useful for	
PT	treating disorders e.g. hypocalcaemia, such as hypocalcaemia, formation	
PT	of calcium stones, mental retardation abnormal body sulfate homeostasis.	
XX		
PS	Disclosure; Page 49-50; Opp; English.	
XX		
CC	The invention relates to 69624 polypeptide, a human transporter family	
CC	member and its corresponding nucleic acid. 69624 protein is useful for	
CC	developing novel diagnostic and therapeutic agents for 69624-mediated or	
CC	related disorders. 69624 protein act as therapeutic or diagnostic agents	
CC	for renal, neurological, colon or hepatic disorders. It act as diagnostic	
CC	targets and therapeutic agents for treating disorders such as	
CC	hypocalcaemia, formation of calcium stones, mental retardation (Canavan	
CC	disease) or abnormal body sulfate homeostasis. 69624 protein may act as	
CC	diagnostic targets and therapeutic agents for controlling cellular	
CC	proliferative and/or differentiative disorders such as carcinoma,	
CC	sarcoma, metastatic disorder or haematopoietic neoplastic disorders e.g.,	
CC	leukaemia, immune disorders such as autoimmune disorders (diabetes	
CC	mellitus, arthritis), multiple sclerosis, encephalomyelitis, myasthenia	
CC	gravis, systemic lupus erythematosus, cardiovascular disorders such as	
CC	arteriosclerosis, atherosclerosis, ischaemia reperfusion injury, cardiac	
CC	hypertrophy, atrial fibrillation etc; and disorders involving abnormal or	
CC	excessive pain. 69624 sequence is useful as pharmacodynamic marker and is	
CC	also used in forensic identification of a biological sample. The present	
CC	sequence is rat sodium-dependent dicarboxylate/sulphate transporter (NaDC	
CC	-1) protein	
XX		
SQ	Sequence 587 AA;	
XX		
Query Match	34.8%; Score 1042.5; DB 7; Length 587;	
Best Local Similarity	38.1%; Pred. No. 6.6e-105;	
Matches 219; Conservative 129; Mismatches 190; Indels 37; Gaps 8;		

QY	26	LVVFLVPLLCPLVMLNEGAFRCMYLLVMAIFWVTEALPLYVTSMPIVAFPIPMGIMS	85
Db	15	LIVLCIPILPLPLIVQTKAYCAYSIIMALLWCTEALPLAVTALFPIVLPFLMGIMD	74
QY	86	SDQTCRLYFKDTLVNFMGGIMVAVENLHKLRLAVIQVCGSPRLHPLHMTWF	145
Db	75	ASEVCIEYFKDTNIFVGLMVAIAVEHNLHKLRLAVIQVCGSPRLHPLHMTWF	134
QY	146	LSMWISNDACTAMMCPILQAVLELQA--QGVCKINHEPOQIV-----GGNK	191
Db	135	LSMWISNTATTAMMVPICHAVLEQGGSKDVEGNNPTFELQECQKEVTKLDNGOP	194
QY	192	KNNEDEP-----PYPTKITLCYLGIAAYASSLGCGTITGATNLTFKGIYEAPRK	242
Db	195	VSAPEPRTKQTEHHRFSQGLSLC-----ICVSASIGGIATLTGTPNLVQGVNSLFP	250
QY	243	NEQMDPPTFMFYSVPSMLVTLTFVFLQWHFGLMRPKSKEAQEQVQREGADVAKK	302
Db	251	QNGNVNFAFWGFAFPTMIILLLAWLQVLFVGNFRKNFGFGEGERKQA--APQ	308
QY	303	VIDQYKDLGPMISHEIQVMILFIMVVMYFTRKPGIFLGWADLLNSKDIRSMPT---	358
Db	309	VIKTQIRLLGPMSPFAEKTIVLVFLVNLVWTFREGFPFGWGDVTFANEKGOSMASDGV	368
QY	359	-IFVVMCFMPLPANYAFRLYCTRGGVPVPTGTPSLITWKFTQTKVPGNLVFLGGFAL	417
Db	369	AIFFISLVMPFIIPSKIPGLMQDPKPKLKA--PPAILTWKTVNDKMPNIVILGGGFAL	426
QY	418	AEQSQSGMAKLIGNALIGLVPLNSVLLVILVAVFLTAFFSSNVAIANIIPVLAEMS	477
Db	427	AKGSEQSGLEWLDKLTPLQHIPPSATAVILCLLAIETECTSNVATTLFPLILASNA	486
QY	478	LAIEIHPYLLILPAGLACSMAFHLPVSTPPNALVAGYANIRTKDMAIAGIGPTIITL	537
Db	487	QAICLHPLVMLPCTLAASLAFMLPVATPPNAIVFSFGGLKVSMDMARAGFLNIIIGLAI	546
QY	538	FVFCQWGLVVPVNLNSPPEWAQIYAAAAAGNKH 572	
Db	547	TLISNSWSIPIF-KLDTFSPWAHSNTSQCLNPSN 580	
RESULT 8			
ID	ADP64799		
XX	ADP64799 standard; protein; 581 AA.		
AC	ADP64799;		
XX			
DT	26-AUG-2004 (first entry)		
XX			
DE	Zebrafish Na+-coupled citrate transporter protein.		
XX			
KW	sodium-coupled citrate transporter; transmembrane citrate transporter;		
KW	lifespan; weight reduction; weight gain prevention; blood cholesterol;		
KW	triglyceride; low density lipopylaaccharide; glucose; obesity;		
KW	hyperlipidemia; hypercholesterolemia; INDY protein.		
XX			
OS	Danio rerio.		
XX			
PN	WO2004048925-A2.		
XX			
PD	10-JUN-2004.		
XX			
PF	20-NOV-2003; 2003WO-US037054.		
XX			
PR	22-NOV-2002; 2002US-0428469P.		
PR	01-APR-2003; 2003US-0459441P.		
XX			
PA	(MEDI-) MEDICAL COLLEGE GEORGIA RES INST.		
PA	(GANR/) GANAPATHY V.		
PA	(INOU/) INOUE K.		
XX	(FEIY/) FEI Y.		
PI	Ganapathy V, Inoue K, Fei Y;		

WPI; 2004-460797/43.
N-PSDB; ADP64798.

New isolated polynucleotide encoding a Na+-coupled citrate transporter (NaCT) polypeptide, useful as a drug target for the treatment of obesity, hyperlipidemia, and hypercholesterolemia.

Claim 42; SEQ ID NO 12; 186pp; English.

The invention relates to novel Na+-coupled citrate transporter proteins and their encoding genes. Inhibitors of transmembrane citrate transporters are useful for extending the lifespan, reducing weight, preventing weight gain or lowering blood cholesterol, triglyceride, LDL or glucose levels in a subject. The NaCT polypeptide is useful as a drug target for the treatment of obesity, hyperlipidemia, and hypercholesterolemia. This sequence corresponds to the zebrafish Na+-coupled citrate transporter protein.

Sequence 581 AA;

Query Match 34.7%; Score 1039; DB 8; Length 581;
Best Local Similarity 38.0%; Pred. No. 1.6e-104;
Matches 221; Conservative 117; Mismatches 185; Indels 58; Gaps 10

Qy 23 WK---GLVVFVLPLCLPVMLLNEGAERFCMYLLLVMAIFWVTEALPLYVTSMIPIVAPP 79
||| :
Db 10 WKMKNTLILFCTPFLLPLPLVIGSKEAGCAVVVLMMAVWCTEVLPAVTALLPAVLFP 69

Qy 80 IMGTMSDDTCRLYPKDFLVMFMGGIMVALAVEYCNLHKRLAARVIQIVGSGPRHLHGL 139
||| :
Db 70 LFRIMEQDVCQMQLKDTNMFLFGGLMAYAVAEHWNLHKRIARLVLLGVRRPALMLGF 129

Qy 140 IMVTMFLSMISNAACATAMCPIIOAVLEELQAQGVCKINHEPOIIVGGNKKNEDRPP 199
||| :
Db 130 MGVTAFLSMWLSNTATTANMPVIVQAVLEQLN-----NTAQEQSSIPETEEKSTEKQPE 184

Qy 200 YPTK-----ITLCYLGIAYASSLGCGGTIIGT 227
||| :
Db 185 SPGEKVVLNGDNFSMESDPDEHSREABERLKWSKGLTLC---VCYAASIGGTATLTGT 240

Qy 228 ATNLTFIGIYEARKNSTEQMDPTTFMFYSVPMLVYTLLTFVFLQWHFMGLWRPKSKEA 287
||| :
Db 241 GPNLVLMGQMSQLFPNDPDIIINFASFGEFAFPNNMIIMTLAWLMIQIVFLGINFKTWGC 300

Qy 288 QEVQRREGADVAKVIDORVDLGPMSIHFIQVMILFIMVVMYFTRKPGIFLGWADLL 347
||| :
Db 301 GTVKTEKEIA--AYNVIKEEHRSLGPMTFGBLSVALFILLVLMWFTDRDPGVGDWATRF 358

Qy 348 NSKD---IRNSMPTIFVVMCMFLPNANYAFRLRYC---TRRGCPVP---TGTPSPSLITWKF 398
||| :
Db 359 FNADKEFVTDATVAVFVAALLFVFPSPKP--RLCFWRTESPTVQOESGPTPALLTWKV 416

Qy 399 IQTKVPGLVFLGGGFALAGSKQSOMAKUIGNALIGLKVLPSNVLVLLVILVAVFUTA 458
||| :
Db 417 TQKWNPWSIIILLGGGFALAGSBTSIGLSKWLGMQSPLOQSIIPWAIATAIVCLMATPTE 476

Qy 459 FSSNVAIANIIPVLAEMSLAIEIHPYLIIPAGLACSMAPHLVPSTPPNALVAGYANIR 518
||| :
Db 477 CTSNVATATFLPLIASMSQSIGVNPLVVMVPCPTLSASFAPMLPVATPPNAIVFSYGLYK 536

Qy 519 TKDMAIAGIPTTIITITLFFVFCQTQGLVYVYNLNSPFEWA 559
||| :
Db 537 VSDMAKTGIWNVIIGILSITLAINSWGRAIF-SLDTFPSWA 576

RESULT 9
ABB82950
ID ABB82950 standard; protein; 592 AA.
XX AC ABB82950;
XX DT 14-APR-2003 (first entry)

XX	Human SLC13A related protein (GenBank Identifier No. GI#2499523) .
DE	SLC13A; p53; sodium-sulfate cotransporter 2; cytosstatic; cancer;
XX	transmembrane protein; human.
KW	Homo sapiens.
OS	WO200298468-A1.
XX	12-DEC-2002.
PN	03-JUN-2002; 2002WO-US017460.
XX	05-JUN-2001; 2001US-0296076P.
PR	10-OCT-2001; 2001US-0328605P.
XX	15-FEB-2002; 2002US-0357253P.
PR	01-MAR-2002; 2002US-0361196P.
XX	(EXEL-) EXELIXIS INC.
PA	Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
XX	Lioubin WN;
PI	WPI; 2003-167297/16.
PL	Identifying candidate p53 pathway modulating agent for diagnosing or
DR	treating cancer comprises detecting test agent-biased activity of an
XX	assay system comprising purified Sodium sulfate cotransporter 2 (SLC13A)
PT	polypeptide.
PT	Claim 13; Page 59-61; 69pp; English.
XX	The invention relates to identifying a candidate p53 pathway modulating
PS	agent. The method involves assaying purified Sodium-sulfate cotransporter
XX	2 (SLC13A) polypeptide or nucleic acid or a functionally active fragment
CC	or derivative with a test agent. The methods are useful for identifying a
CC	candidate p53 pathway modulating agent, modulating a p53 pathway of a
CC	cell, or a mammalian cell and for diagnosing a disease in a patient. The
CC	disease is breast, colon, lung or ovary cancer having greater than 25%
CC	expression level. The method is useful for manufacturing a medicament for
CC	diagnosing or treating breast, colon, lung or ovary cancer. Sequences
CC	AB882949-952 represent polypeptide sequences related to the human SLC13A
CC	protein
XX	Sequence 592 AA;
SQ	Query Match 34.3%; Score 1026; DB 6; Length 592;
	Best Local Similarity 37.2%; Pred. No. 4.4e-103;
	Matches 219; Conservative 124; Mismatches 189; Indels 56; Gaps 10;
Qy	26 LVVFLVLLCLPVMLNEGAEFRCMYLLLVMAIFMVTALPLYTSMIPVAFFPMGIMS 85
Dd	15 LIVFVFVILLPULPVPSKEYACAYAIILMALFWCTEAPLAVTAUPLFLPFMMGLVD 74
Qy	86 SDOTCRLYFKDTLVFMFGGMVALAVBYCNLHKRLALRVIQVGCSRRRLFGLIMTMVF 145
Dd	75 ASEVAVEYLKDSNLLFFGGLVALAVEHNLHKRIALRVLLIVGRVAPLLILGFMVLTFAP 134
Qy	146 LSMWISNACTAMWCPIQAVLEBELOAGVCCKINHEFOYQIVGNKKNNDEPPYPVK-- 203
Dd	135 LSMWISNTATSAMWVPFAHAVLDQLHSQSANNVEE-----GSNNPTFELQEPSPOKEV 188
Qy	204 -----ITLCYYLGIA YASSLGGCGTIGTATNLTPKG 235
Dd	189 TKLDNGQALPVTSSASSEGRAHLKSQHLLHTOCHSLCVYSASIGGIATLTGTANLVLQG 248
Qy	236 IYEARFNKNSTEQMDFPFTMFYSVPMSLVYTLLTFVFLQHFMGIL-WRPKSKAEAEVQRGR 294
Dd	249 QINSLFQONGNVNFASFWSFAPTWTVILLALLAWLWLQIILFLGFNFNRKFCIGEKMQ--- 305
Qy	295 EGADVAKKVIDORYKDLGPMSIHIEIQMWILFI FMVMVYTRKPGIFLGWADLL--NSKD- 351

RESULT 9
ABB82950
ID ABB82950 standard; protein; 592 AA.
XX
XX ABB82950;
XX
DT 14-APR-2003 (first entry)

Db 306 EQQAAVCVITQTEHRLGPMTEFAEKASILEFVLVLLWFTTRBEGFGLGNLAFPNKGE 365
QY 352 --TRNSMPTIFVVMCMFLPANYAFRLRYCTRRGGPV--PTGPTPSLITWKFIQTKVPWGL 407
Db 366 SMVSDGTVAIFIGIIMFIISKPGLTQDPENFGKLPGLG---LLDWTNQNKPWNI 421
QY 408 VFLGGGFALAEKSGKSGMAKLIGNALIGKVLPNVLLVLLVILVAVLTFAPSSNVAIAN 467
Db 422 VLLGGGYALAKGSGSLSEWLGNKLTPLQSVPAFAIAIILSLVATFECTSNVATT 481
QY 468 IIIIVLAEMSLAIEIHPYLLIPLAGLACSMAPHLPVSTPPNALVAGYANIRTKDMAIAGI 527
Db 482 IFLPILASMAQAICLHPLVYMLPCTLATSIAFMLPVATPNAIVFSGDLKVLDMARAGF 541
QY 528 GPTIITIILFVFCQWGLVVPNLNSFPPEWQAIIYAAA----ALGNKT 571
Db 542 LLNIIGVLIILALAINSGIPLF-SLHSFSPSAQSNNTTAQCLPSLANTT 588

RESULT 10
ABB82951
ID ABB82951 standard; protein; 592 AA.
XX AC ABB82951;
XX DT 14-APR-2003 (first entry)
XX DE Human SLC13A related protein (GenBank Identifier No. GI#4506979).
XX KW SLC13A; p53; sodium-sulfate cotransporter 2; cytostatic; cancer;
XX KW transmembrane protein; human.
XX OS Homo sapiens.
XX PN WO200298468-A1.
XX PD 12-DEC-2002.
XX PF 03-JUN-2002; 2002WO-US017460.
XX PR 05-JUN-2001; 2001US-0296076P.
XX PR 10-OCT-2001; 2001US-0328605P.
XX PR 15-FEB-2002; 2002US-0357253P.
XX PR 01-MAR-2002; 2002US-0361196P.
XX PA (EXEL-) EXELIXIS INC.
XX PI Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
XX PI Lioubin MN;
XX DR WPI; 2003-167297/16.
XX DR N-PSDB; AB224204.
XX PT Identifying candidate p53 pathway modulating agent for diagnosing or
XX PT treating cancer comprises detecting test agent-biased activity of an
XX PT assay system comprising purified Sodium sulfate cotransporter 2 (SLC13A)
XX PT polypeptide.
XX PS Claim 13; Page 62-64; 69pp; English.
XX CC The invention relates to identifying a candidate p53 pathway modulating
XX CC agent. The method involves assaying purified Sodium-sulfate cotransporter
XX CC 2 (SLC13A) polypeptide or nucleic acid or a functionally active fragment
XX CC or derivative with a test agent. The methods are useful for identifying a
XX CC candidate p53 pathway modulating agent, modulating a p53 pathway of a
XX CC cell, or a mammalian cell and for diagnosing a disease in a patient. The
XX CC disease is breast, colon, lung or ovary cancer having greater than 25%
XX CC expression level. The method is useful for manufacturing a medicament for
XX CC diagnosing or treating breast, colon, lung or ovary cancer. Sequences
XX CC ABB82949-952 represent polypeptide sequences related to the human SLC13A
XX CC protein
XX SQ Sequence 592 AA;

Query Match 34.3%; Score 1026; DB 6; Length 592;
Best Local Similarity 37.2%; Pred. No. 4.e-103; Indels 56; Gaps 10;
Matches 219; Conservative 124; Mismatches 189;
QY 26 LVVFLVPLLCPLPMLNNEGABFRCCMYLLLVMAIFWVVTALPLYVTSMPITVAFPPMGIMS 85
Db 15 LIVFFVPLLLPLPLTVPSKEAYCAIILMALFWCTEALPLAVTALPLILFPMGIVD 74
QY 86 SDQTRLVFKDTLVNFMGGIWMALAVEYCNLHKRLALRVIQIVGCSPRRLHFLIMVTFM 145
Db 75 ASEVAEVELKDSNLLFFGGLVAIAVEHNLKRLALRVLLIVGVRRPAPLILGLFVLVTF 134
QY 146 LSMWISNACTAMMCPITIOAVLEELQAQGVCKINHEPOQIVGVGNKKNEDPPPTK-- 203
Db 135 LSMWISNATATSMWVPIAHAVLDQLHSSQASSNVSE-----GSNNPTFELQPSPOKEV 188
QY 204 -----ITLCYYIGIAYASSLGGCGTIIIGTATNLTFKG 235
Db 189 TKLDNGQALPVTSASSEGRAHLSQKHLHLTQCMSLCVCYSASIGGIATLTGTAPNLVQG 248
QY 236 IYEAFKNSQTEQMDPTTFMFYSVPNMLVYTLTLTFVLOWHFMGL-WRPKSKEAQEVQGR 294
Db 249 QINSUFPQNGVNVNFASFSPFAPPTNVILLLANLWLQILFLGFNFRKNFGIGKQW--- 305
QY 295 EGADVAKKVIDORYKDLGPMSTHEIQVMILFIEMVVMYFTRKPGIFLGWADLL--NSKD- 351
Db 306 EQQAAVCVITQTEHRLGPMTEFAEKASILEFVLVLLWFTTRBEGFGLGNLAFPNKGE 365
QY 352 --IRNSMPTIFVVMCMFLPANYAFRLRYCTRRGGPV--PTGPTPSLITWKFIQTKVPWGL 407
Db 366 SMVSDGTVAIFIGIIMFIISKPGLTQDPENFGKLPGLG---LLDWTNQNKPWNI 421
QY 408 VFLGGGFALAEKSGKSGMAKLIGNALIGKVLPNVLLVILVAVLTFAPSSNVAIAN 467
Db 422 VLLGGGYALAKGSGSLSEWLGNKLTPLQSVPAFAIAIILSLVATFECTSNVATT 481
QY 468 IIIIVLAEMSLAIEIHPYLLIPLAGLACSMAPHLPVSTPPNALVAGYANIRTKDMAIAGI 527
Db 482 IFLPILASMAQAICLHPLVYMLPCTLATSIAFMLPVATPNAIVFSGDLKVLDMARAGF 541
QY 528 GPTIITIILFVFCQWGLVVPNLNSFPPEWQAIIYAAA----ALGNKT 571
Db 542 LLNIIGVLIILALAINSGIPLF-SLHSFSPSAQSNNTTAQCLPSLANTT 588

RESULT 11
ABR40100
ID ABR40100 standard; protein; 592 AA.
XX AC ABR40100;
XX DT 24-JUL-2003 (first entry)
XX DE Human renal sodium dicarboxylate co-transporter NADCL.
XX DE Anti-diabetic; anorectic; sodium dicarboxylate cotransporter; human;
XX KW sodium tricarboxylate cotransporter; HepNadC; hepatocyte; HepG2;
XX KW diabetes; obesity; lipid metabolism; aging; NADCL.
XX OS Homo sapiens.
XX PN WO2003029465-A1.
XX PD 10-APR-2003.
XX PF 27-SEP-2002; 2002WO-JP010038.
XX PR 28-SEP-2001; 2001JP-00299433.
XX PR 28-AUG-2002; 2002JP-00249016.
XX PA (SAKA) OTSUKA PHARM CO LTD.
XX SQ

Kanemoto N, Omori Y, Sugano S, Obuchi Y;
WPI; 2003-354728/33.

Sodium di- or tricarboxylate cotransporter gene (*HepNADC*) for treatment
and prevention of diabetes, obesity, for improving lipid metabolism and
aging.

Example 1; Fig 1; 75pp; Japanese.

The present invention relates to human sodium di- or tricarboxylate
cotransporter protein (*HepNADC*; ABR40097). *HepNADC* gene is expressed in
human hepatocyte cancer-origin cells HepG2. Compounds which control the
expression of the *HepNADC* gene and activity of the *HepNADC* polypeptide
may be useful for treating and preventing diabetes, obesity, for
improving lipid metabolism and aging. The present sequence is the *NADCI*
protein (GenBank U26209), which was used in a sequence alignment with the
HepNADC sequence

Sequence 592 AA;

```
Query Match      34.3%; Score 1026; DB 6; Length 592;
Best Local Similarity 37.2%; Pred No. 4.4e-103;
Matches 219; Conservative 124; Mismatches 189; Indels 56; Gaps 10

QY   26 LVVFLVLCLLPVMNLNEGAFRCMYLLLVNAIFVTAEALPLYTSMIPIVAFPINGMS 85
    :|::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
DB   15 LIVFFVFILLPLPILVPKSEAYCAYAIIIMALFMCTEALPLAVTALPLLILFPMMGIVD 74

QY   86 SDQTCLRYPFDITLWNFGGMIVALAVEYNLUHLRALRVIOVGSPRRLLHFLIMTMP 145
    :|::|||::|||::|||::|||::|||::|||::|||::|||::|||::
DB   75 ASEVAVEYLKDSNLIFFGLGLVALTAVERNHLKRIALRVLIIIIGVRPALPLLGFMLVATFA 134

QY   146 LSMWSINAACTAMMCPIIQAIVLELOAGGVCKINHPEYOIVGGNKKNDEPPVPTK-- 203
    |||||||:|||::|||::|||::|||::|||::|||::|||::|||::|||::
DB   135 LSMWSINTAISMAMPVIHAHVLDQLHSQSANNVE-----GNPNTPFEIQESPKQEV 188

QY   204 -----ITLCYYILGIAYASSLGCSCGTIIIGTANLTPKG 235
        ::|||::|||::|||::|||::|||::|||::|||::|||::|||::
DB   189 TKLDNGQALPVTSASSSRAHLSQKLHLTLQCMSLCVCYSASIGGIATLTGTAPNLVLQG 248

QY   236 IYEARFNKSNTQMDFPFMFYSVPSMLVYTLLTVFWLQHFMGL-WRPKSKEAQEQVQR 294
    :|::|||::|||::|||::|||::|||::|||::|||::|||::|||::
DB   249 QINSFPQNGNVNFASFWSFPAFTMVIIALLAWLWLQILFLGFNRXNCFGIKEMQQ--- 305

QY   295 EGADVAKKVIDORYKDLAGPMSIHIEIQMWILFIWMVVYFKPGIFELGWDL--NSKD- 351
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
DB   306 EQOAAVCYVQTEHRLLGPMTFAEKASILFVLVLWFTRHPGFFLWGNNLATFNAKGE 365

QY   352 --IRNSSPTIFFVVMCEMLPANFYARYCTRRTGGPV--PTGPTPSLTIWKETIQTKVPMGL 407
    :|::|||::|||::|||::|||::|||::|||::|||::|||::|||::
DB   366 SMSWDGTVAFIGIIMIIFIISKFPLTGDPENPGKKAPLG----LLDKWTVNQMPWNI 421

QY   408 VFLLGGFALAEGSKSGMAKLIGNALI GLVKLPNSVLLLVLVILVAVFLTAFSSNVAIAN 467
    :|::|||::|||::|||::|||::|||::|||::|||::|||::|||::
DB   422 VLLGGGYALAKGSERSGLSEWLGNKLTPLQSVAPAIAIITLSLLVATFTECTSNVATT 481

QY   468 IIIPVLAEMSIAIEIHPYLILPAGLACSMAPHLPVSSTPNNALVAGYANIETKMATAGI 527
    :|::|||::|||::|||::|||::|||::|||::|||::|||::|||::
DB   482 IFLPILASMAQAACLHPLYVNWLPCTLATSALFAMPLPATPPPNAIVFSFGDLKVLDMWRAGR 541

QY   528 GPRIITIITLFVFCOTGVLGVVYNLNISFPWEAAQTYAAA----ALGNKT 571
    |::|||::|||::|||::|||::|||::|||::|||::|||::|||::
DB   542 LNMII GVLIITALAINSWGIPLP-SLHFSPSWAQSNTTAACQLPSLANYT 598
```

RESULT 12
ABRS7025
ID ABRS7025 standard; protein; 592 AA.
XX AC ABRS7025;
XX DT 05-AUG-2003 (first entry)

XX	Human NADc-1 amino acid sequence.
DE	
XX	
KW	Human; TCHI69; dicarboxylate transport; hepatotropic; cytostatic;
KW	nephrotropic; vasotropic; antidiabetic; liver disease; hepatitis;
KW	hepatic sclerosis; alcohol-related liver disease; prostate disease;
KW	prostatitis; prostatic hypertrophy; spleen disease; spleen hyperactivity;
KW	kidney disease; nephritis; kidney failure; nephritis; dropsy; diabetes;
KW	diabetes-associated renal disease; metabolic disease; hyperlipaemia;
KW	circulatory disease; arteriosclerosis; cancer; NADc-1.
OS	Homo sapiens.
XX	
PN	WO2003025168-A1.
XX	
PD	27-MAR-2003.
XX	
PF	13-SEP-2002; 2002WO-JP009444.
XX	
PR	17-SEP-2001; 2001JP-00281992.
PR	02-OCT-2001; 2001JP-00306873.
PR	16-APR-2002; 2002JP-00113279.
XX	
PA	(TAKE) TAKEDA CHEM IND LTD.
XX	
PI	Nakanishi A, Uno Y, Sagiya Y;
DR	WPI; 2003-313352/30.
XX	
PT	Protein TCHI69 with dicarboxylate transport activity for treatment and
PT	diagnosis of diseases including liver disease, cancer and circulatory
PT	disorders.
XX	
PS	Example 1; Fig 1-2; 132pp; Japanese.
XX	
CC	The present invention describes protein TCHI69 and its salts having
CC	dicarboxylate transport activity. TCHI69 has hepatotropic, cytostatic,
CC	nephrotropic, vasotropic and antidiabetic activities. The TCHI69 protein
CC	and polynucleotide can be used in the treatment, prevention and diagnosis
CC	of liver disease (such as hepatitis, hepatic sclerosis and alcohol-
CC	related liver disease); prostate disease (such as prostatitis and
CC	prostatic hypertrophy); spleen disease (such as spleen hyperactivity);
CC	kidney disease (such as nephritis, kidney failure, nephritis, dropsy and
CC	diabetes-associated renal disease); metabolic disease (such as diabetes);
CC	circulatory disease (such as hyperlipaemia and arteriosclerosis); and
CC	cancer (such as non-small cell lung cancer, liver cancer, renal cancer,
CC	ovarian cancer, prostate cancer, stomach cancer, pancreatic cancer,
CC	breast cancer, colon cancer, bladder cancer and womb cancer). The present
CC	invention represents a human NADc-1 amino acid sequence which is given in
CC	comparison with the human TCHI69 amino acid sequence in an example from
XX	the present invention
SQ	Sequence 592 AA;
	Query Match 34.3%; Score 1026; DB 6; Length 592;
	Best Local Similarity 37.2%; Pred. No. 4.4e-103;
	Matches 219; Conservative 124; Mismatches 199; Indels 56; Gaps 10
QY	26 LVVFLVPLLCLPVMNLNEGAEFRCMYLLLVAFWVTETALPLVTSMPIVATPINGMS 85
DB	15 LVVFVFPILLPLTLPVDSKEAYCAYAIILMALFWCTETALPLAVTALFPLLFPFMGIVD 74
QY	86 SDQTCRLYFKDTLWMFGIIMVALAVEYNLHKRLALRVLIQVGCSPPRLHFLGLIMVTWF 145
DB	75 AGEVAVEYLKDSNLLFFGLGLVALAVSEWNLHKRIALRVLLIIGVPAPAILGLFGMLVTAF 134
QY	146 LGMWTISNAECTAMMCPIIQAVLEELQAQGCKINHPEQYQIVGGNKKNNEDEPPYPTEK-- 203
DB	135 LSNMTSNTATSAMVVPITAHVLVDQLHSQSASSNVEE-----GSNNPTTELQEPSQKEV 188
QY	204 -----ITLCYYLGIAVASSLGCGGTITIGTATNTLTFKG 235
DB	189 TKLDNGQALPVTSASSSGRAHLKSQHKLHLTKCMSLCVCYSASICGATATGTAPNLVLQGS 248

RESULT 12
ABR57025
ID ABR5
XX
AC ABR5
XX
DT 05-A

QY 236 IYERARFKNSTEQMDPFTFMFYSVPMSLVYTLTFVFLQWHFMGL-WRPKSKEAQEVQR 294
Db 249 QINSLFPQNGVNVNFASFWSFAPFTTWVILLALLAWLWLQILFLGFNFRKNFGIGERKQ--- 305
QY 295 EGADVAKKVDQRYKDLGPMSEIHEIQVMILFIFMVMYFTRKPGIFLGWADLL--NSKD- 351
Db 306 EQQAAAYCVIQTETRLHGLGPMTFEAKAISILFVLVLLWFTREPGFGLGNGLAFPPNAKGE 365
QY 352 --IRNSMPTTFVVMCFMPLPANYAFRLYCTRRGGPV--PTGPTPSLITWKFIQTKVPWGL 407
Db 366 SMVSDGTVAIFIGIIMFIIPSKPGLTQDPENPGKLKAPLG---LLDWTNVQKMPWNI 421
QY 408 VFLGGGFALAEBSKQSGMAKLIGNALIGKVLPSNVLLVILVAVFLTAFSSNVAIAN 467
Db 422 VLLGGGYALAKGERSGLSEWLNKLTPLQSPVAPAPAIILSLLVATFTECTSNVATT 481
QY 468 IIPVLAEMSLATEIHPYLVLPCTLATSALFMLPVATPPNAIVFSGDLKVLDMARAGF 541
Db 482 IFDILASMAQAICLHPLYVNLPCLTATSLAFMLPVATPPNAIVFSGDLKVLDMARAGF 541
QY 528 GPTIITITLTFVFCQTVGVLVYVNLNSFPPEWQAIIYAAA----ALGNKT 571
Db 542 LLNIIGVLIATLAINSGIPLF-SLHSPFSPSAQSNNTTAQCLPLSLANTT 588

RESULT 13
ADE63755
ID ADE63755 standard; protein, 592 AA.
XX AC ADE63755;
XX DT 29-JAN-2004 (first entry)
XX DE Human Protein Q13183, SEQ ID NO 9699.
XX KW Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX OS Homo sapiens.
XX PN WO2003016475-A2.
XX XX 27-FEB-2003.
XX PF 14-AUG-2002; 2002WO-US025765.
XX PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX (GEO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX PI Woolf C, D'urso D, Befort K, Costigan M;
XX WPI; 2003-268312/26.
DR GENBANK; Q13183.
XX PT New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX Claim 1; Page; 1017pp; English.

CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence

CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. Gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX Sequence 592 AA;
SQ

Query Match 34.3%; Score 1026; DB 7; Length 592;
Best Local Similarity 37.2%; Pred. No. 4.4e-103;
Matches 219; Conservative 124; Mismatches 189; Indels 56; Gaps 10;
QY 26 LVVELVPLCLCPVMLINEGAEFRMYLLLVMAIFWVTEALPLYVTSMIPIVAPPIMGIMS 85
Db 15 LIVFFVPLILLPLILVPSKEAYCAIILMALFWCTEALPLAVTALFLLIPMMGIVD 74
QY 86 SDQTCRLYFKDTLVMPMGIMVALAVEYCNLHKRLALRVQIVGCSPPRIHLGIMWTF 145
Db 75 ASEVAVEYLKDSNLLFFGGLLVAVAEHWNHKLALRVLLIVGVRPAPLILGFMVLVTA 134
QY 146 LSMWISNAACTAMCPDIOAVLEELQAGVCKINHEPQYIVGNGKNNEDEBPYPTK-- 203
Db 135 LSMWISNTATSAMVPIAHAVLDQJHSSQASSNVEE-----GSNNPTTELQEPSQKEV 198
QY 204 -----ITLCYVLGIAYASSLGGCGTIIIGTATNLTFFKG 235
Db 189 TKLDNGQALPVTSASSEGRAHLKQKHLTLTQCMSCVCSASIGGIATLTGTAPNLVLOG 248
QY 236 IYERARFKNSTEQMDPFTFMFYSVPMSLVYTLTFVFLQWHFMGL-WRPKSKEAQEVQR 294
Db 249 QINSLFPQNGVNVNFASFWSFAPFTTWVILLALLAWLWLQILFLGFNFRKNFGIGERKQ--- 305
QY 295 EGADVAKKVDQRYKDLGPMSEIHEIQVMILFIFMVMYFTRKPGIFLGWADLL--NSKD- 351
Db 306 EQQAAAYCVIQTETRLHGLGPMTFEAKAISILFVLVLLWFTREPGFGLGNGLAFPPNAKGE 365
QY 352 --IRNSMPTTFVVMCFMPLPANYAFRLYCTRRGGPV--PTGPTPSLITWKFIQTKVPWGL 407
Db 366 SMVSDGTVAIFIGIIMFIIPSKPGLTQDPENPGKLKAPLG---LLDWTNVQKMPWNI 421
QY 408 VFLGGGFALAEBSKQSGMAKLIGNALIGKVLPSNVLLVILVAVFLTAFSSNVAIAN 467
Db 422 VLLGGGYALAKGERSGLSEWLNKLTPLQSPVAPAPAIILSLLVATFTECTSNVATT 481
QY 468 IIPVLAEMSLATEIHPYLVLPCTLATSALFMLPVATPPNAIVFSGDLKVLDMARAGF 541
Db 482 IFDILASMAQAICLHPLYVNLPCLTATSLAFMLPVATPPNAIVFSGDLKVLDMARAGF 541
QY 528 GPTIITITLTFVFCQTVGVLVYVNLNSFPPEWQAIIYAAA----ALGNKT 571
Db 542 LLNIIGVLIATLAINSGIPLF-SLHSPFSPSAQSNNTTAQCLPLSLANTT 588

RESULT 14
ADG16995
ID ADG16995 standard; protein, 619 AA.
XX AC ADG16995;
XX XX

PR 12-JUN-2002; 2002US-0387934P.
 PR 12-JUN-2002; 2002US-0387960P.
 PR 12-JUN-2002; 2002US-0388022P.
 PR 12-JUN-2002; 2002US-0388096P.
 PR 12-JUN-2002; 2002US-0388432P.
 PR 12-JUN-2002; 2002US-0388479P.
 PR 13-JUN-2002; 2002US-0389123P.
 PR 14-JUN-2002; 2002US-0389120P.
 PR 14-JUN-2002; 2002US-0389146P.
 PR 17-JUN-2002; 2002US-0389742P.
 PR 18-JUN-2002; 2002US-0389604P.
 PR 18-JUN-2002; 2002US-0389884P.
 PR 19-JUN-2002; 2002US-0390066P.
 PR 19-JUN-2002; 2002US-0390144P.
 PR 19-JUN-2002; 2002US-0390209P.
 PR 25-JUN-2002; 2002US-0391726P.
 PR 06-AUG-2002; 2002US-0401628P.
 PR 09-AUG-2002; 2002US-0402688P.
 PR 12-AUG-2002; 2002US-0402822P.
 PR 13-AUG-2002; 2002US-0403458P.
 PR 15-AUG-2002; 2002US-0403617P.
 PR 15-AUG-2002; 2002US-0403732P.
 PR 26-AUG-2002; 2002US-0406182P.
 PR 12-SEP-2002; 2002US-0410085P.
 PR 13-SEP-2002; 2002US-0410505P.
 PR 23-SEP-2002; 2002US-0412955P.
 PR 30-SEP-2002; 2002US-0415195P.
 PR 23-OCT-2002; 2002US-0420627P.
 PR 23-OCT-2002; 2002US-0420718P.
 PR 24-OCT-2002; 2002US-0420852P.
 PR 31-OCT-2002; 2002US-0422750P.
 PR 01-NOV-2002; 2002US-0423095P.
 PR 05-NOV-2002; 2002US-0423748P.
 (CURA-) CURAGEN CORP.
 XX
 PI Alsobrook JP, Anderson DW, Baumgartner JC, Berghs C, Boldog FL;
 PI Burgess CE, Caeman SJ, Catterton E, Dhanabal M, Edinger SR;
 PI Ellerman K, Ettenberg S, Gangolli EA, Gerlach VL, Gorman L;
 PI Grosse WM, Gunther E, Guo X, Gusev VV, Herrmann JL, Ji W, Kekuda R;
 PI Kramasov NV, Larochelle WJ, Li L, Liang H, Low K, MacDougall JR;
 PI MacIsachian T, Malyokhar M, McQueney K, Mezick AJ, Miller CS;
 PI Millet I, Padigar M, Patturajan M, Peyman JA, Qian X, Rastelli L;
 PI Rieger DK, Rothenberg ME, Shenoy SG, Shimkets RA, Smithson G;
 PI Spytek KA, Stone DJ, Sukumaran S, Szekeres ES, Vernet CAM, Voss EZ;
 PI Wolenc AR, Zhong H, Zhong H;
 XX
 DR WPI; 2004-053467/05.
 DR N-PSDB; ADH42442.
 XX
 PT New NOVX polypeptides and nucleic acid molecules useful for preventing or
 PT treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
 PT atherosclerosis or diabetes, in chromosome mapping, tissue typing or in
 PT pharmacogenomics.
 XX
 PS Claim 2; SEQ ID NO 996; 1503pp; English.
 XX
 CC The invention relates to 566 new isolated human polypeptides and their
 CC encoding genes, sequences that are at least 95% identical to these or
 CC sequences comprising one or more conservative substitutions in these. The
 CC polypeptide, polynucleotide and antibodies against the polypeptides are
 CC useful in diagnosing, treating or preventing NOVX-associated disorders,
 CC e.g. cardiomyopathy, atherosclerosis, hypertension, cancer, obesity,
 CC diabetes, AIDS, multiple sclerosis, graft-versus-host disease,
 CC Alzheimer's disease, Parkinson's disease, asthma, or fertility disorders.
 CC The nucleic acids are further used as hybridization probes, in chromosome
 CC mapping, tissue typing, preventive medicine, and pharmacogenomics. The
 CC polypeptides are also useful as vaccines. This sequence represents an
 CC example of the polypeptide of the invention.
 XX
 SQ Sequence 576 AA;

Query Match

33.6%; Score 1005; DB 8; Length 576;

Best Local Similarity 36.8%; Pred. No. 8.5e-101;
 Matches 214; Conservative 123; Mismatches 198; Indels 46; Gaps 11;
 QY 11 PPVKCSNFFANHWKGLAVV-FLVPLLLCLPVMLLNEGAFFRCMYLLLVMAIFWVTEALPLYV 69
 DB 4 PTMASALSVSXPKSPFVILFVTPLLPLVILMPAKFVRCAYVYIILMAIYWCTEVIPLAV 63
 QY 70 TSMIPIVAFPIMGIMSSDOTCRLYFKDTLVMFMGGMVALAVEYCNLHKLRLALRVQIVG 129
 DB 64 TSLMPVLLFPLQLDSRQVCVQYMKDTNMLFGLGLIVAVAVERNLHKLRIALTLLWVG 123
 QY 130 CSPRLHFLGLIMVTFMFLSMWISNAACTAMWCPPIQAVLEELQAQVC-----KI 178
 DB 124 AKPARLMLFGMGTALLSMWISNTATTAMVPIVEAILQOMEATSATAEAGLELVKGA 183
 QY 179 NHEPQYQIVGNKNKNEDEPPYTKITLCYLLGIAYASSILGGCGTIICTATNLTFFGIYE 238
 DB 184 KELPGSQVIFEGFTLGGQEDQERKRLCKAMTLCIYAASIGGTATLTGTGPNVVLGQMN 243
 QY 239 ARPKNSTEQMDPFTFMFYSPVPSMLVYTLTLTFVFLQWHEM-----GLWRPKSKEAQ 288
 DB 244 ELFPDSKDLVNFASWFAFAFPNMLVMLLFAWLWLOFVYRNFNFKSWGCGLESKKNEKA- 302
 QY 289 EVQREGADVAKVJDQRYKDLGPMHSIHEIQMILFIFWVVMYFTRKPGIFLGWADLL- 347
 DB 303 -----ALKVQLQEEYRKGLPLSPFAEINVLICFFLLVILMFSRDPGFMGWLTVAV 351
 QY 348 ---NSKDIRNSMPTIFVVMVCFMPLPA---NYAFRLYCTRRGGPVPTGPTSLITWKFIQT 401
 DB 352 VEGETKYVSDATVAIFVATLLFIVPSQPKFNF-RSQTEERKTPFPY-PPLLDWKVTQE 409
 QY 402 KVPWGLVFLGGGFALAEKSGKQMAKLTGNALIGKLPNSVLLIIV-ILVAVFLTAFS 460
 DB 410 KVPWGLVLLGGGFALAKGSEASGLSVWNGKQMEPLHAVPPAAITLISLLVAVF-TECT 468
 QY 461 SNVAIANIIPVLAEMSLAIEIHPYLIILPAGLACSMAPHLPVSTPPNALVAGYANIRTK 520
 DB 469 SNVATTLFLPIFASMSRSIGLNPLYIMLPCTLSASFAPMLPVATPPNAIVFTYGLKVA 528
 QY 521 DMAIAGIGPTIITITLTFVFCOTGVLVYVYVPLNSRPEWAQI 561
 DB 529 DMWKTGVIMNIIGVFCVFLAVNTWGRAIF-DLDHFPDWNANV 568

Search completed: June 30, 2005, 08:52:01
 Job time : 85.7044 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2005, 08:36:57; Search time 21.9924 Seconds
(without alignments)
1941.542 Million cell updates/sec

Title: US-10-017-479A-2
Perfect score: 2994
Sequence: 1 MBIEIGEQPPVKCSNFPA.....NSFPWAQIYAAALGNKTH 572

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/1aa/5A-COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B-COMB.pep:*
3: /cgn2_6/prodata/1/1aa/6A-COMB.pep:*
4: /cgn2_6/prodata/1/1aa/6B-COMB.pep:*
5: /cgn2_6/prodata/1/1aa/PCTUS-COMB.pep:*
6: /cgn2_6/prodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1742	58.2	368	4	US-09-270-767-42029
2	1026	34.3	599	4	US-09-949-016-9866
3	898	30.0	169	4	US-09-270-767-57286
4	896	29.9	626	4	US-09-556-916-20
5	895.5	29.9	627	4	US-09-556-916-8
6	891	29.8	626	4	US-09-556-916-14
7	891	29.8	626	4	US-09-556-916-22
8	890.5	29.7	627	4	US-09-556-916-2
9	889.5	29.7	627	4	US-09-556-916-10
10	889.5	29.7	627	4	US-09-949-016-6840
11	886	29.6	626	4	US-09-556-916-16
12	885.5	29.6	627	4	US-09-556-916-4
13	879.5	29.4	601	4	US-09-949-016-9977
14	760	25.4	580	4	US-09-556-916-24
15	759.5	25.4	581	4	US-09-556-916-12
16	755	25.2	580	4	US-09-556-916-18
17	754.5	25.2	581	4	US-09-556-916-6
18	711	23.7	561	4	US-09-949-016-8161
19	605.5	20.2	527	4	US-09-602-787A-516
20	577	19.3	230	4	US-09-270-767-43713
21	498.5	16.6	180	4	US-09-270-767-59102
22	463	15.5	194	4	US-09-270-767-42669
23	457	15.3	470	4	US-09-543-681A-5952
24	397	13.3	233	4	US-09-489-847-176
25	362	12.1	132	4	US-09-270-767-57987
26	357	11.9	335	4	US-09-602-787A-518
27	344	11.5	524	3	US-09-134-001C-5457

28	258.5	8.6	421	4	US-09-248-796A-20749	Sequence 20749, A
29	187.5	6.3	624	4	US-09-543-681A-4343	Sequence 4343, Ap
30	179.5	6.0	548	4	US-09-902-540-11870	Sequence 11870, A
31	176.5	5.9	596	4	US-09-902-540-13547	Sequence 13547, A
32	172.5	5.8	493	4	US-09-540-236-2120	Sequence 2120, Ap
33	166.5	5.6	302	4	US-09-902-540-10445	Sequence 10445, A
34	162	5.4	470	4	US-09-438-185A-209	Sequence 209, App
35	159	5.3	430	3	US-09-134-001C-2981	Sequence 2981, Ap
36	150.5	5.0	547	4	US-09-489-039A-13843	Sequence 13843, A
37	147	4.9	443	4	US-09-602-787A-532	Sequence 532, App
38	137	4.6	363	4	US-09-248-796A-20740	Sequence 20740, A
39	137	4.6	363	4	US-09-270-767-42613	Sequence 42613, A
40	136.5	4.6	478	3	US-09-134-001C-4637	Sequence 4637, Ap
41	136.5	4.6	494	4	US-09-543-681A-7033	Sequence 7033, Ap
42	135	4.5	694	4	US-09-252-991A-22637	Sequence 22637, A
43	134.5	4.5	436	4	US-09-583-110-4729	Sequence 4729, Ap
44	134.5	4.5	440	4	US-09-107-433-5201	Sequence 5201, Ap
45	134	4.5	437	4	US-09-543-681A-6984	Sequence 6984, Ap
46	131.5	4.4	614	4	US-09-489-039A-12605	Sequence 12605, A
47	131	4.4	319	4	US-09-270-767-57927	Sequence 57927, A
48	123	4.1	449	4	US-09-328-352-7512	Sequence 7512, Ap
49	122	4.1	465	4	US-09-543-681A-5337	Sequence 5337, Ap
50	120	4.0	569	2	US-08-750-723A-2	Sequence 2, Appli
51	120	4.0	569	3	US-09-191-275-2	Sequence 2, Appli
52	116.5	3.9	323	4	US-09-489-039A-7408	Sequence 7408, Ap
53	115.5	3.9	370	4	US-09-902-540-10305	Sequence 10305, A
54	114.5	3.8	502	4	US-09-252-991A-23327	Sequence 23327, A
55	112	3.7	477	3	US-09-134-001C-3487	Sequence 3487, Ap
56	111	3.7	457	4	US-09-489-039A-12762	Sequence 12762, A
57	111	3.7	484	4	US-09-107-532A-3830	Sequence 3830, Ap
58	110	3.7	429	4	US-08-311-731A-287	Sequence 287, App
59	107	3.6	476	4	US-09-543-681A-5260	Sequence 5260, Ap
60	106	3.5	494	2	US-09-031-392-5	Sequence 5, Appli
61	106	3.5	494	3	US-09-299-549-5	Sequence 5, Appli
62	106	3.5	494	3	US-08-610-417-5	Sequence 5, Appli
63	105	3.5	446	4	US-09-252-991A-17185	Sequence 17185, A
64	105	3.5	451	4	US-09-328-352-5922	Sequence 5922, Ap
65	104	3.5	367	3	US-09-134-001C-5557	Sequence 5557, Ap
66	104	3.5	398	4	US-09-489-039A-13159	Sequence 13159, A
67	103.5	3.5	381	4	US-09-710-279-3284	Sequence 3284, Ap
68	103.5	3.5	394	4	US-09-134-000C-3732	Sequence 3732, Ap
69	102	3.4	116	4	US-09-602-787A-520	Sequence 520, App
70	102	3.4	478	4	US-09-134-000C-6219	Sequence 6219, Ap
71	102	3.4	487	4	US-09-489-039A-8887	Sequence 8887, Ap
72	102	3.4	527	4	US-09-252-991A-18859	Sequence 18859, A
73	102	3.4	1014	4	US-09-252-991A-29868	Sequence 29868, A
74	101.5	3.4	503	4	US-09-583-110-4277	Sequence 4277, Ap
75	101.5	3.4	536	4	US-09-107-433-2945	Sequence 2945, Ap
76	101	3.4	321	4	US-09-107-532A-4118	Sequence 4118, Ap
77	100	3.3	345	4	US-09-902-540-9836	Sequence 9836, Ap
78	99	3.3	211	4	US-09-198-452A-1166	Sequence 1166, Ap
79	99	3.3	349	4	US-09-902-540-13275	Sequence 13275, A
80	98.5	3.3	441	4	US-09-489-039A-10612	Sequence 10612, A
81	98	3.3	557	4	US-09-540-236-2206	Sequence 2206, Ap
82	97.5	3.3	444	4	US-09-603-208A-240	Sequence 240, App
83	96.5	3.2	447	3	US-09-134-001C-4523	Sequence 4523, Ap
84	96.5	3.2	474	4	US-09-489-039A-14147	Sequence 14147, A
85	96.5	3.2	630	3	US-09-134-001C-4615	Sequence 4615, Ap
86	96	3.2	514	4	US-09-252-991A-30952	Sequence 30952, A
87	95.5	3.2	273	4	US-09-042-709A-19	Sequence 19, Appli
88	95	3.2	279	4	US-09-134-000C-5300	Sequence 5300, Ap
89	95	3.2	466	4	US-09-438-185A-231	Sequence 231, App
90	95	3.2	511	4	US-09-679-686B-12	Sequence 12, Appli
91	95	3.2	713	4	US-09-543-681A-6094	Sequence 6094, Ap
92	94.5	3.2	404	4	US-09-328-352-6361	Sequence 6361, Ap
93	94.5	3.2	574	2	US-08-948-569A-4	Sequence 4, Appli
94	94.5	3.2	574	2	US-09-188-469-4	Sequence 4, Appli
95	94.5	3.2	574	3	US-09-397-238A-4	Sequence 4, Appli
96	94.5	3.2	635	2	US-09-014-969-11	Sequence 11, Appli
97	94	3.1	462	2	US-08-898-976-2	Sequence 2, Appli
98	94	3.1	462	2	US-08-898-976-4	Sequence 4, Appli
99	94	3.1	471	4	US-09-543-681A-5580	Sequence 5580, Ap
100	94	3.1	502	4	US-09-902-540-13393	Sequence 13393, A

ALIGNMENTS

RESULT 1

US-09-270-767-42029
; Sequence 42029, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 42029
; LENGTH: 368
; TYPE: PRT
; ORGANISM: *Drosophila melanogaster*
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-42029

Query Match 58.2%; Score 1742; DB 4; Length 368;
Best Local Similarity 93.8%; Pred. No. 9.4e-180;
Matches 333; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 1 MEIIGQPPQPVKCSNFFAHNMKGLVFLVPLLCLEPVMILNEGAFCRCMYLLVMAIFW 60
Db 14 MEIIGQPPQPVKCSNFFAHNMKGLVFLVPLLCLEPVMILNEGAFCRCMYLLVMAIFW 73

Qy 61 VTEALPVTSMIPDIVAPPINGMSSDQTLXPKDTLVNFMGIMVALAVEYCNLHKL 120
Db 74 VTEALPVTSMIPDIVAPPINGMSSDQTLXPKDTLVNFMGIMVALAVEYCNLHKL 133

Qy 121 ALRVIVGCSPRHLFGLIMVTMFLSNWISNAACTAMMCPHIOAVLEELQAQGVCKINH 180
Db 134 ALRVIXXXXXXXXXXXXXXXXXXWISNAACTAMMCPHIOAVLEELQAQGVCKINH 193

Qy 181 EPQYQIVGGKNNKNEDEPPYPTKITLCYILGIAYASSLGGCGTIIIGTATNLTFKGIYEAR 240
Db 194 EPQYQIVGGKNNKNEDEPPYPTKITLCYILGIAYASSLGGCGTIIIGTATNLTFKGIYEAR 253

Qy 241 FKNSTEQMDPPTFMFYSVPVSMVLTLLTFVFLQWHFMGLWRPKSKEAQEVQRGEGADVA 300
Db 254 FKNSTEQMDPPTFMFYSVPVSMVLTLLTFVFLQWHFMGLWRPKSKEAQEVQRGEGADVA 313

Qy 301 KKVIDORYKDLGPMISHEIQVMILFIFMVVYTRKPGIFLGWADLNSKDIRNS 355
Db 314 KKVIDORYKDLGPMISHEIQVMILFIFMVVYTRKPGIFLGWADLNSKDIRNS 368

RESULT 2

US-09-949-016-9866
; Sequence 9866, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 9866
; LENGTH: 599
; TYPE: PRT
; ORGANISM: Human
; US-09-949-016-9866

Query Match 34.3%; Score 1026; DB 4; Length 599;
Best Local Similarity 37.2%; Pred. No. 6.9e-102;
Matches 219; Conservative 124; Mismatches 189; Indels 56; Gaps 10;

Qy 26 LVVFLVPLLCLEPVMILNEGAFCRCMYLLVMAIFWVTEALPVTSMIPDIVAPPINGMIS 85
Db 22 LIVFFVPIILLPLILVPSKEAYCAVAILMALFWCTEALPLAVTALFFLILFPMGIVD 81

Qy 86 SDQTCRLYKDTLVNFMGIMVALAVEYCNLHKLRLRVLIQIVGCSPRHLFGLIMVTMF 145
Db 82 ASEVAVEYLKDSNLLFFGGLLVAVAEHWNHKLRLRVLIQIVGCSPRHLFGLIMVTMF 141

Qy 146 LSNWISNAACTAMMCPHIOAVLEELQAQGVCKINHBPQYQIVGGKNNKNEDEPPYPTK-- 203
Db 142 LSNWISNATATSMVPIAHAVLDQLHSSQASSNVBE-----GSNNPTFELQEPSQKEV 195

Qy 204 -----ITLCYILGIAYASSLGGCGTIIIGTATNLTFKGI 235
Db 196 TKLDNGQALPVTSSASSEGRAHLSQKHLTLTQCMSLCVCYSASIGGTATLTGTAPNLVLOG 255

Qy 236 IYEAREFKNSTEQMDPPTFMFYSVPVSMVLTLLTFVFLQWHFMGLWRPKSKEAQEVQRGR 294
Db 256 QINSLEFPQNGVNVNFASFWSFAFTWVILLALLAWLWQLILFLGFNFRKNFGIGEKMQ--- 312

Qy 295 EGADVAKKVIDORYKDLGPMISHEIQVMILFIFMVVYTRKPGIFLGWADLL--NSKD- 351
Db 313 EQQAAAYCVIQTREHRLGPMTEAKAISILFVLVLLWFTREBPGFLGWNLAFFNAKGE 372

Qy 352 --IRNSMPTIFVVMCMFPLPANYAFILRYCTRRGPPV--PTGPTPSLITWKFIQTKVPWGL 407
Db 373 SMVSDGTVAIFGIIMFIIPSFGPLTQDPENPGKUKAPLG----LDDWRTVNNQKMPWNI 428

Qy 408 VFLLGGGFALEAGSKOSGMAKLIGNALIGLVLPNSVLLLVILVAVFLTAFSSNVAIAN 467
Db 429 VLLGGYALAKGSESGSLSEWLNKLTPLQSVPAIPAIAIILSLVATPTECTSNVATTT 488

Qy 468 IIPVLAESLAIEIHPYLIILPAGLACSMAPHLPVSTPPNALVAGYANIRTKDMAIAGI 527
Db 489 IFLPILASMAQAICLHPLVYMLPCTLATSLAFMLPVATPPNAIVFSFGDLKVLDMARAGF 548

Qy 528 GPTIIITLITLVFCQTWGLVYVYENLNSFPEWAOIYAAA----ALGNKT 571
Db 549 LLMNIIGVLIILALAINSWGIPLP--SLHSFFSWAOSNTTAQCLPSLANTT 595

RESULT 3

US-09-270-767-57286
; Sequence 57286, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 57286
; LENGTH: 169
; TYPE: PRT
; ORGANISM: *Drosophila melanogaster*
US-09-270-767-57286

Query Match 30.0%; Score 898; DB 4; Length 169;
Best Local Similarity 100.0%; Pred. No. 7.2e-89;
Matches 169; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

187	Qy	VGNKKNNDEPPYPTKIITLCYYLGIAYASSLGGCGTIIIGTATNLT	246
1	Db	VGNKKNNDEPPYPTKIITLCYYLGIAYASSLGGCGTIIIGTATNLT	60
247	Qy	QMDPPTFMFYSVPMSLVYTLTFVFLQWHPNGLMRPKSKEAQEVQR	306
61	Db	QMDPPTFMFYSVPMSLVYTLTFVFLQWHPNGLMRPKSKEAQEVQR	120
307	Qy	RYKDLGPMGSHIEIQWMLIFFMVWYMYFTRKPGIFLGWADLLNSK	355
121	Db	RYKDLGPMGSHIEIQWMLIFFMVWYMYFTRKPGIFLGWADLLNSK	169

RESULT 4

US-09-556-916-20
; Sequence 20, Application US/09556916
; Patent No. 6548271

; FACILE NO. 6548271
: GENERAL INFORMATION:

```

; GENERAL INFORMATION:
; APPLICANT: Turner, Alex
; APPLICANT: zambrowicz, Brian
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins
; FILE REFERENCE: 8535-041-999
; CURRENT APPLICATION NUMBER: US/09/556,916
; CURRENT FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 20
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-556-916-20

```

Query Match 29.9%; Score 896; DB 4; Length 626;
Best Local Similarity 32.5%; Pred. No. 8.9e-88;
Matches 205; Conservative 122; Mismatches 188. Indels 1

[illegible]

Db	184	KNSQPSLELIFVNEESNADLTTLTHNENLNGVPSITNPIKTANQHQGKQHPQSQEPQVL	243
Qy	199	-VPPTK-----ITLCVYLGIAYASSLGGCGTIIIGTATNLTFGKGYEARFKN	243
Db	244	TSPRKQKLNRYSRHHDQICKLUSLSISVSATIGLTTIIGTSLSL-----IFLEHFN	299
Qy	244	---STEQMDPTFMFYSPSLVLTTLTFVLEWHFGL-WRPKSEAEQVQRREGADV	299
Db	300	QYPAAEVNVFGTWLFPSPISLIMLVSWFMHWLFLGCNFKETCSLSKKKTKRE-Q	357
Qy	300	AKKVIDORYKDLGPMSEIHEIOMVILTFPMVVMYTRKPGILGWADLLNSKDIR-NSMPT	358
Db	358	SEKRIQEYKEKLGDISYEMWTGFFIIMTLVLTREPFGVPGWGDGFEEKKGYRTDATVS	417
Qy	359	IFVVMCFMLPANYAFLRYCTRRCGSPVPTG-----TPSLITWKFIOKVPWGLVFL	411
Db	418	VFLGFLFLIIPAK----KPCF---GKNKDGENQHSLSGTESIITWKDPQKTHPEWELIVL	470
Qy	412	GGGFALAEBSKOSGMAKLIGNALGLKVLPN-SVLLLVILVAVFLTAFSSNVAIANII	470

Db	471	GGGYALASGSKSGLSTWIGNQMLSSLSLPWAVTLACILVSI-VTEPVSNDPATITIFL	529
Qy	471	PVLAEMSLAIEIHPLYLILPAGLACSMAPHLPVSTPPNALVAGYANIRTKDWAIGPT	530
Dd	530	PILCSLSETLHINPLTYLIPVTMCISFAVMLPVGNPNNAIVFSYGHQCQIKDMVKAGLVN	589
Qy	531	IITIITLFVFCQTGLGVVYPNLNSFPPEAAQI	561
Dd	590	VIGLVIWMVAINTGWSLFL-HLDYTPAARV	619

RESULT 5

RESULTS 3
US-09-556-916-8
; Sequence 8, Application US/09556916

; Patent No. 6548271

```
, RECORD NUMBER: 8621
, GENERAL INFORMATION:
, APPLICANT: Turner, Alex
, APPLICANT: Zambrowicz, Brian
, APPLICANT: Nehls, Michael
, APPLICANT: Friedrich, Glenn
, APPLICANT: Sands, Arthur T.
, TITLE OF INVENTION: No. 6548271el Human Transporter Proteins
, FILE REFERENCE: 8535-041-999
, CURRENT APPLICATION NUMBER: US/09/556,916
, CURRENT FILING DATE: 2000-04-21
, NUMBER OF SEQ ID NOS: 32
, SOFTWARE: FastSeq for Windows Version 3.0
, SEQ ID NO 8
, LENGTH: 627
, TYPE: PRT
, ORGANISM: Homo sapiens
US-09-556-916-8
```

Query Match 29.9%; Score 895.5; DB 4; Length 627;
Best Local Similarity 32.4%; Pred. No. 1e-87;
Matches 205; Conservative 122; Mismatches 188; Indels 117

Qy	24	KGLVVFVLPLCLPVMLLNNEGAEFRCMYLLLVMAIFWYTEALPLVVTSMIPVAFFPMGI	83
		: :	
Dd	12	KLLLVVCVPLLPLPVLPSPSEASCAVLIIVTAVVYVSEAVPLGAAALVPAPLAFPPFGV	71
		: :	
Qy	84	MSDQTCRLFYKDTLVNFMGMIMVALVEYNLHKRIALRVIOVGCSRRRHFLGLIWMV	143
		: :	
Dd	72	LRSNEVAAYEYFKNTLLLVGVICVAAAVEKWNLHKRIALRMVLNAGAKPGMLLLCFMCC	131
		: :	
Qy	144	MEFSMWISNAACTAMMCPIIQAVLEELQAQGCKVINHEPOYQIVGGNKKNNEDEP	198
		: :	
Dd	132	TLLSMWLISNTSTTANVMPIVEAVLOEL-----VSAEDE-OLVAGNSNTEABEPILDV	183
		: :	
Qy	199	-----	198

[illegible]

QY 470 IPVLAEMSLAIEIHPYLLIPAGLACSMAPHLPVSTPPNALVAGYANIRTKDMAIAGIP 529
Db 530 LPILCSLSETHINPLTYLIPVTCISFAVMLPVGNPPNAIVFSYGHQCIKDMVKAGLV 589
QY 530 TIITITLTFVFCOTWGLVVPNLNSRPEWAQI 561
Db 590 NVIGLVIMVAINTWGVSLF-HLDTYPAWARV 620

RESULT 6
US-09-556-916-14
; Sequence 14, Application US/09556916
; Patent No. 6548271
; GENERAL INFORMATION:
; APPLICANT: Turner, Alex
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins
; FILE REFERENCE: 8535-041-999
; CURRENT APPLICATION NUMBER: US/09/556,916
; CURRENT FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-556-916-14

Query Match 29.8%; Score 891; DB 4; Length 626;
Best Local Similarity 32.3%; Pred. No. 3.1e-87;
Matches 204; Conservative 122; Mismatches 189; Indels 116; Gaps 15;

QY 24 KGLVVLVPLCLPVMMLNEGAEFRCMYLLVMAIFWVTEALPLYVTSMIPVAFPIMG 83
Db 12 KLLLVVCPVLLLPVLHPSSSEASCAYLVITAVVWSEAVPLGAALVPAFLYPPFGV 71
QY 84 MSSDOTCLRYFKDTLVMPMGIMVALAVEYNLHKLRLALRVIOVGCSPRRHLHGLIMVT 143
Db 72 LRSNEVAAYEFKNTLLLVGVICVAAAVEKNLHKLRLALRVLMVLMAGAKPGMLLCFCMCT 131
QY 144 MFLSMWISNACTAMMCPICQVLEELQAGVCKINHEPOQIVGGNKKNEDEP----- 198
Db 132 TLLSMWLSNTSTTAMVMPIVEAVLOEL-----VSAEDE-QLVAGNSNTEEAEPISLDV 183
QY 199 ----- 198
Db 184 KNSQPSLELIFVNEESNADLTTLMHENLNGVPSITNPIKTANQHQKQHPQKQV 243
QY 199 -PYPTK-----ITLCYVLGIAYASSLGGCGGTIICTATNLTFKGIYEAFKN 243
Db 244 TPSPRKQKLNRYSHHDDMICKLSLSISYSATIGLTIIGTSTL-----IFLEHFN 299
QY 244 ---STEQMDPTFMFYSPSMVYLLTFVFLQWHEMGL-WRPKSKAEQVQREGADV 299
Db 300 QYPAAEVNFGTFLFSPISLIMLVSWFWHMLFLGCNFKETCSLSKKKTKRE--QL 357
QY 300 AKKVIDORYKDLGPMSEIHEIQVMILFIFVMVYFTRKPGIFLGWADLLNSKDRI-NSMPT 358
Db 358 SEKRIOEYKLGDISYPEWVTGFFILMTVLTFTREPFGVPGWDSFPEKKGYRTDVS 417
QY 359 IFVVVMCFMLPANYAFIRYCTRGGPVPTGP-----TPSLITWKFIQTKVPKGLVFL 411
Db 418 VFLGFLFLIPAK-----KPCF---GKNDGENQEHSLGTEPIITWKDFQKTMPEIVILV 470
QY 412 GGGFALAEKSGKSMAGLIGNALIGLVLPN-SVLLLVILVAVFLTAFSSNVAIANII 470
Db 471 GGGYALASGKSGSLGTWIGNQMLSSLPWAVTLLACILVSI-VTEFVSNPATITIFL 529
QY 471 PVLAEMSLAIEIHPYLLIPAGLACSMAPHLPVSTPPNALVAGYANIRTKDMAIAGIPT 530

Db 530 PILCSLSETHINPLTYLIPVTCISFAVMLPVGNPPNAIVFSYGHQCIKDMVKAGLVN 589
QY 531 ITITITLTFVFCOTWGLVVPNLNSRPEWAQI 561
Db 590 VIIGLVIMVAINTWGVSLF-HLDTYPAWARV 619

RESULT 7
US-09-556-916-22
; Sequence 22, Application US/09556916
; Patent No. 6548271
; GENERAL INFORMATION:
; APPLICANT: Turner, Alex
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins
; FILE REFERENCE: 8535-041-999
; CURRENT APPLICATION NUMBER: US/09/556,916
; CURRENT FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 22
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-556-916-22

Query Match 29.8%; Score 891; DB 4; Length 626;
Best Local Similarity 32.5%; Pred. No. 3.1e-87;
Matches 205; Conservative 121; Mismatches 189; Indels 116; Gaps 15;

QY 24 KGLVVLVPLCLPVMMLNEGAEFRCMYLLVMAIFWVTEALPLYVTSMIPVAFPIMG 83
Db 12 KLLLVVCPVLLLPVLHPSSSEASCAYLVITAVVWSEAVPLGAALVPAFLYPPFGV 71
QY 84 MSSDOTCLRYFKDTLVMPMGIMVALAVEYNLHKLRLALRVIOVGCSPRRHLHGLIMVT 143
Db 72 LRSNEVAAYEFKNTLLLVGVICVAAAVEKNLHKLRLALRVLMVLMAGAKPGMLLCFCMCT 131
QY 144 MFLSMWISNACTAMMCPICQVLEELQAGVCKINHEPOQIVGGNKKNEDEP----- 198
Db 132 TLLSMWLSNTSTTAMVMPIVEAVLOEL-----VSAEDE-QLVAGNSNTEEAEPISLDV 193
QY 199 ----- 198
Db 184 KNSQPSLELIFVNEESNADLTTLMHENLNGVPSITNPIKTANQHQKQHPQKQV 243
QY 199 -PYPTK-----ITLCYVLGIAYASSLGGCGGTIICTATNLTFKGIYEAFKN 243
Db 244 TPSPRKQKLNRYSHHDDMICKLSLSISYSATIGLTIIGTSTL-----IFLEHFN 299
QY 244 ---STEQMDPTFMFYSPSMVYLLTFVFLQWHEMGL-WRPKSKAEQVQREGADV 299
Db 300 QYPAAEVNFGTFLFSPISLIMLVSWFWHMLFLGCNFKETCSLSKKKTKRE--QL 357
QY 300 AKKVIDORYKDLGPMSEIHEIQVMILFIFVMVYFTRKPGIFLGWADLLNSKDRI-NSMPT 358
Db 358 SEKRIOEYKLGDISYPEWVTGFFILMTVLTFTREPFGVPGWDSFPEKKGYRTDVS 417
QY 359 IFVVVMCFMLPANYAFIRYCTRGGPVPTGP-----TPSLITWKFIQTKVPKGLVFL 411
Db 418 VFLGFLFLIPAK-----KPCF---GKNDGENQEHSLGTEPIITWKDFQKTMPEIVILV 470
QY 412 GGGFALAEKSGKSMAGLIGNALIGLVLPN-SVLLLVILVAVFLTAFSSNVAIANII 470
Db 471 GGGYALASGKSGSLGTWIGNQMLSSLPWAVTLLACILVSI-VTEFVSNPATITIFL 529
QY 471 PVLAEMSLAIEIHPYLLIPAGLACSMAPHLPVSTPPNALVAGYANIRTKDMAIAGIPT 530
Db 530 PILCSLSETHINPLTYLIPVTCISFAVMLPVGNPPNAIVFSYGHQCIKDMVKAGLVN 589

QY 531 IITITITLVFCQWGLVVPVNLNSPFEWAQI 561
Db 590 VIGLVIVMVAINTWGVSLF-HLDTYPAMARV 619

RESULT 8

US-09-556-916-2
; Sequence 2, Application US/09556916
; Patent No. 6548271
; GENERAL INFORMATION:
; APPLICANT: Turner, Alex
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins
; FILE REFERENCE: 8535-041-999
; CURRENT APPLICATION NUMBER: US/09/556,916
; CURRENT FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 627
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-556-916-2

Query Match 29.7%; Score 890.5; DB 4; Length 627;
Best Local Similarity 32.3%; Pred. No. 3.5e-87;
Matches 204; Conservative 122; Mismatches 189; Indels 117; Gaps 15;
QY 24 KGLVVFLVPLLCPLVMLNEGAFRCMYLLLVMAIFVWTEALPLVVTSMIPVAFPIIMGI 83
Db 12 KLLLVVCPVLLLLPLVPLHPSSEASCAVYLVITAVVWVSEAVPLGAAALVPAFLVPPFGV 71
QY 84 MSSDOTCLRYFKDTLVFMFGGITVVALAVEYCNLHKRLALRVQIVGCSPPRLHFGLIWVT 143
Db 72 LRSNEVAEYFKNTLLLVGVICVAAAVEKNLHKRIALRMVLMAGAKPGMLLLCFMCT 131
QY 144 MFLSNWISNAACTAMCPITIOAVLEELQAQGVCKINHPQYQIVGGNKKNEDEP- 198
Db 132 TLLSNWLSNTSTAMVPIVEAVLQEL- - - - -VSAEDE-QLVAGNSNTEAEPISLDV 183
QY 199 - - - - - 198
Db 184 KNSQPSLELIFVNEDRSNADLTLLMHNENLNGVPSITNPIKTANQHQKQHPQV 243
QY 199 - - - - -PYTK- - - - -ITLCYVIGIAYASSLGGCGTIIIGTATNLTFKGIYEAPFK 242
Db 244 LTPSPRKQKLNKRYSHHDQMICLSLSISYSATIGGLTIIIGTSTSL- - - - -IFLEHN 299
QY 243 N- - - - -STEOMDPPTFMFYSPVSMVYLLTFVFLQWHEMGL-WRPKSKEAQEVQREGAD 298
Db 300 NOYPAAEVNFGTWFLEFPISLIMLVSWFWMHMLFLGCNFKETCSLSKSKKTKRE- - - - -Q 357
QY 299 VAKKVIDORYKDLGPMSEIHEIQVMILFIMVVMYFTRKPGIFLGWADLLNSKDIR-NSMP 357
Db 358 LSEKIQEYKLGDISYPEVMYTGFFILMTVLMFTREPFGVPGWDSFFEKKGYRTDATV 417
QY 358 TIFVVMCFMPLPANYAFIRYCTRGGVPVTPG- - - - -TPSLITWKFIQTKVPMGLVFL 410
Db 418 SVFLGFLFLIPAK- - - - -KPCF- - - - -GKNDGENQHSLSGTPEITWKDFQKTMPEWIVL 470
QY 411 LGGCFALAGSKQSGMAKLGALIGLVLPN-SVLLLVILVAVFLTAFNSNAIANII 469
Db 471 VGGYALAGSKSGSLTWIGNQMLSSLPWAVTLACILVSI-VTEFVSNPATITIF 529
QY 470 IPVLAEMSLAIEIHPYLLILPAGLACSMAFHLPVSTPPNALVAGYANIRTKDMAIAGIP 529
Db 530 LPILCSLSETHINPLTYLIPVTWCISFAVMLPVGNPNNAIVFSYGHQCIQKMWKAGLV 589
QY 530 TIITITITLVFCQWGLVVPVNLNSPFEWAQI 561
Db 590 VIGLVIVMVAINTWGVSLF-HLDTYPAMARV 619

Db 590 NVIGLVIVMVAINTWGVSLF-HLDTYPAMARV 620

RESULT 9

US-09-556-916-10
; Sequence 10, Application US/09556916
; Patent No. 6548271
; GENERAL INFORMATION:
; APPLICANT: Turner, Alex
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins
; FILE REFERENCE: 8535-041-999
; CURRENT APPLICATION NUMBER: US/09/556,916
; CURRENT FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 627
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-556-916-10

Query Match 29.7%; Score 890.5; DB 4; Length 627;
Best Local Similarity 32.4%; Pred. No. 3.5e-87;
Matches 205; Conservative 121; Mismatches 189; Indels 117; Gaps 15;
QY 24 KGLVVFLVPLLCPLVMLNEGAFRCMYLLLVMAIFVWTEALPLVVTSMIPVAFPIIMGI 83
Db 12 KLLLVVCPVLLLLPLVPLHPSSEASCAVYLVITAVVWVSEAVPLGAAALVPAFLVPPFGV 71
QY 84 MSSDOTCLRYFKDTLVFMFGGITVVALAVEYCNLHKRLALRVQIVGCSPPRLHFGLIWVT 143
Db 72 LRSNEVAEYFKNTLLLVGVICVAAAVEKNLHKRIALRMVLMAGAKPGMLLLCFMCT 131
QY 144 MFLSNWISNAACTAMCPITIOAVLEELQAQGVCKINHPQYQIVGGNKKNEDEP- 198
Db 132 TLLSNWLSNTSTAMVPIVEAVLQEL- - - - -VSAEDE-QLVAGNSNTEAEPISLDV 183
QY 199 - - - - - 198
Db 184 KNSQPSLELIFVNEDRSNADLTLLMHNENLNGVPSITNPIKTANQHQKQHPQV 243
QY 199 - - - - -PYTK- - - - -ITLCYVIGIAYASSLGGCGTIIIGTATNLTFKGIYEAPFK 242
Db 244 LTPSPRKQKLNKRYSHHDQMICLSLSISYSATIGGLTIIIGTSTSL- - - - -IFLEHN 299
QY 243 N- - - - -STEOMDPPTFMFYSPVSMVYLLTFVFLQWHEMGL-WRPKSKEAQEVQREGAD 298
Db 300 NOYPAAEVNFGTWFLEFPISLIMLVSWFWMHMLFLGCNFKETCSLSKSKKTKRE- - - - -Q 357
QY 299 VAKKVIDORYKDLGPMSEIHEIQVMILFIMVVMYFTRKPGIFLGWADLLNSKDIR-NSMP 357
Db 358 LSEKIQEYKLGDISYPEVMYTGFFILMTVLMFTREPFGVPGWDSFFEKKGYRTDATV 417
QY 358 TIFVVMCFMPLPANYAFIRYCTRGGVPVTPG- - - - -TPSLITWKFIQTKVPMGLVFL 410
Db 418 SVFLGFLFLIPAK- - - - -KPCF- - - - -GKNDGENQHSLSGTPEITWKDFQKTMPEWIVL 470
QY 411 LGGCFALAGSKQSGMAKLGALIGLVLPN-SVLLLVILVAVFLTAFNSNAIANII 469
Db 471 VGGYALAGSKSGSLTWIGNQMLSSLPWAVTLACILVSI-VTEFVSNPATITIF 529
QY 470 IPVLAEMSLAIEIHPYLLILPAGLACSMAFHLPVSTPPNALVAGYANIRTKDMAIAGIP 529
Db 530 LPILCSLSETHINPLTYLIPVTWCISFAVMLPVGNPNNAIVFSYGHQCIQKMWKAGLV 589
QY 530 TIITITITLVFCQWGLVVPVNLNSPFEWAQI 561
Db 590 NVIGLVIVMVAINTWGVSLF-HLDTYPAMARV 620

```
RESULT 10
US-09-949-016-6840
; Sequence 6840, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6840
; LENGTH: 627
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6840

Query Match      29.7%; Score 889.5; DB 4; Length 627;
Best Local Similarity 32.3%; Pred. No. 4.5e-87;
Matches 204; Conservative 122; Mismatches 189; Indels 117; Gaps 15;

QY 24 KGLVFLVPLLCPLVMLNEGAFRCMYLLLVMAIFWVTEALPLVYVTSMPIVAFPIMG I 83
DB 12 KLLLVVCVPLLLPLVLPVHPSSEASCAYVLIVTAVVWSEAVPLGAALVPAFLYFFGV 71
QY 84 MSSDQTCRLYFKDTLVNFMGIMVALAVECNLHKLRLALRVIQVGCSPRRLLHFLIMVT 143
DB 72 LRSNEVAAYEFKNTLLLVGVICVAAAVEKNWLNHKLRLALRVLWLMAGAKPGMLLLCFMCC 131
QY 144 MFLSMWISNAACTAMMCPPIQAVLEELQAGVCKINHEPOYQIVGKNKNEDEP----- 198
DB 72 LRSNEVAAYEFKNTLLLVGVICVAAAVEKNWLNHKLRLALRVLWLMAGAKPGMLLLCFMCC 131
QY 132 TLLSMWLSNSTTAVMPIVEAVLQEL-----VSAEDE-QLVAGNSNTEEAEPISLDV 183
DB 199 ----- 198
QY 184 KNSQPSLELIFVNEEDRSNADLTTLHMNENLNGVPSITNPIKTANQHQKQHPQKQV 243
DB 199 --PYPTK-----ITLCYVLGIAYASSLGGCGTIIGTATNLTFGIYEAFK 242
DB 244 LTPSPRKQKLNRYRSHHDQMICKLSLSISYSATIGLTLTIGTSTL-----IFLHFN 299
QY 243 N---STEQMDPPTFMFYSVPSMLVYLLTFVFLQWHEMGL-WRPKSKAEQVQREGAD 298
DB 300 NQYPAAEVNVNFGTWFLFSFPLSLIMLVVSWFWMHFLGCFKETSLSKKKTKRE--Q 357
QY 299 VAKVIOQRYKDLGPMSTHEIQVMILFIWVMVYFTRKPGIFLGWADLLNSKDIR-NSMP 357
DB 358 LSEKRIQEEYKLGDISYPMVMTGFFILMTVLVFTREPFGVPCWDSFFEKKGYRTDATV 417
QY 358 TIFVVMCFMPLPANYAFLRYCTRGGVPVTPG-----TPSLITWKPIQTKVPWGLVFL 410
DB 418 SVFLGFLFLIIPAK-----KPCF---GKKNNGENQEHSLGTEPIITWKDFQKTMPEVIL 470
QY 411 LGGFFALASGSKSGMAKLGALNIGLKVLPN-SVLLLVLLVAVLTAFTSSNVAIANII 469
DB 471 VGGYALASGSKSGSLTWIGNQMLSSLPWPAVTLTLLACILVSI-VTEFVSNPATITIF 529
QY 470 IPVLAESLAIEIHPYLILPAGLACSMAFHLVSTTPPNALVAGYANIRTKDMAIAGIGP 529
DB 530 LPILCSLSETHNPLTYLIPVTCISFVAVMLPVGNPNNAIVFSYGHQCIKDMVKAGLV 589
QY 530 TIIITILFVPCQTWGLVVPVNLNSFPPEWAI 561
DB 590 NVIGLVIMVAINTGVSLF-HLDYTPAWARV 620
```

```
RESULT 11
US-09-556-916-16
; Sequence 16, Application US/09556916
; Patent No. 6548271
; GENERAL INFORMATION:
; APPLICANT: Turner, Alex
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; TITLE OF INVENTION: No. 6548271 Human Transporter Proteins
; FILE REFERENCE: 8535-041-999
; CURRENT APPLICATION NUMBER: US/09/556,916
; CURRENT FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-556-916-16
```

```
Query Match      29.6%; Score 886; DB 4; Length 626;
Best Local Similarity 32.3%; Pred. No. 1.1e-86;
Matches 204; Conservative 121; Mismatches 190; Indels 116; Gaps 15;

QY 24 KGLVFLVPLLCPLVMLNEGAFRCMYLLLVMAIFWVTEALPLVYVTSMPIVAFPIMG I 83
DB 12 KLLLVVCVPLLLPLVLPVHPSSEASCAYVLIVTAVVWSEAVPLGAALVPAFLYFFGV 71
QY 84 MSSDQTCRLYFKDTLVNFMGIMVALAVECNLHKLRLALRVIQVGCSPRRLLHFLIMVT 143
DB 72 LRSNEVAAYEFKNTLLLVGVICVAAAVEKNWLNHKLRLALRVLWLMAGAKPGMLLLCFMCC 131
QY 144 MFLSMWISNAACTAMMCPPIQAVLEELQAGVCKINHEPOYQIVGKNKNEDEP----- 198
DB 132 TLLSMWLSNSTTAVMPIVEAVLQEL-----VSAEDE-QLVAGNSNTEEAEPISLDV 183
QY 199 ----- 198
DB 184 KNSQPSLELIFVNEESNADLTTLHMNENLNGVPSITNPIKTANQHQKQHPQKQV 243
QY 199 --PYPTK-----ITLCYVLGIAYASSLGGCGTIIGTATNLTFGIYEAFK 243
DB 244 TSPSPRKQKLNRYRSHHDQMICKLSLSISYSATIGLTLTIGTSTL-----IFLHFN 299
QY 244 ---STEQMDPPTFMFYSVPSMLVYLLTFVFLQWHEMGL-WRPKSKAEQVQREGADV 299
DB 300 QYPAAEVNVNFGTWFLFSFPLSLIMLVVSWFWMHFLGCFKETSLSKKKTKRE--QL 357
QY 300 AKVIOQRYKDLGPMSTHEIQVMILFIWVMVYFTRKPGIFLGWADLLNSKDIR-NSMPT 358
DB 358 SEKRIQEEYKLGDISYPMVMTGFFILMTVLVFTREPFGVPCWDSFFEKKGYRTDATV 417
QY 359 IFVVMCFMPLPANYAFLRYCTRGGVPVTPG-----TPSLITWKPIQTKVPWGLVFL 411
DB 418 VFLGFLFLIIPAK-----KPCF---GKKNNGENQEHSLGTEPIITWKDFQKTMPEVILV 470
QY 412 GGGFALASGSKSGMAKLGALNIGLKVLPN-SVLLLVLLVAVLTAFTSSNVAIANII 470
DB 471 GGGYALASGSKSGSLTWIGNQMLSSLPWPAVTLTLLACILVSI-VTEFVSNPATITIFL 529
QY 471 PVLAESLAIEIHPYLILPAGLACSMAFHLVSTTPPNALVAGYANIRTKDMAIAGIGP 530
DB 530 PILCSLSETHNPLTYLIPVTCISFVAVMLPVGNPNNAIVFSYGHQCIKDMVKAGLVN 589
QY 531 IITITILFVPCQTWGLVVPVNLNSFPPEWAI 561
DB 590 VIGLVIMVAINTGVSLF-HLDYTPAWARV 619
```


85 261 8.7 162 11 US-09-833-245-1948 Sequence 1948, Ap
86 260.5 8.7 163 15 US-10-424-599-184881 Sequence 184881,
87 258 8.6 236 16 US-10-767-701-43427 Sequence 43427, A
88 257.5 8.6 162 11 US-09-833-245-1945 Sequence 1945, Ap
89 241 8.0 259 16 US-10-739-930-9458 Sequence 9458, Ap
90 234 7.8 226 15 US-10-296-115-980 Sequence 980, App
91 232.5 7.8 300 16 US-10-767-701-46317 Sequence 46317, A
92 228.5 7.6 174 16 US-10-425-115-335649 Sequence 335649,
93 228 7.6 78 9 US-09-864-761-34252 Sequence 34252, A
94 228 7.6 78 14 US-10-029-386-28038 Sequence 28038, A
95 222 7.4 457 15 US-10-369-493-4266 Sequence 4266, Ap
96 217 7.2 135 15 US-10-276-774-1592 Sequence 1592, Ap
97 211 7.0 164 15 US-10-424-599-205152 Sequence 205152,
98 185 6.2 482 15 US-10-335-977-5881 Sequence 5881, Ap
99 184 6.1 487 15 US-10-369-493-23559 Sequence 23559, A
100 179 6.0 494 15 US-10-335-977-5882 Sequence 5882, Ap

ALIGNMENTS

RESULT 1

US-10-167-994-2
; Sequence 2, Application US/10167994
; Publication No. US20030082647A1
; GENERAL INFORMATION:
; APPLICANT: Reenan, Robert A.
; APPLICANT: Rogina, Blanka
; APPLICANT: Helfand, Stephen L.
; TITLE OF INVENTION: TRANSPORTER PROTEIN
; FILE REFERENCE: 13407-013001
; CURRENT APPLICATION NUMBER: US/10/167,994
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/255,013
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 10/017,479
; PRIOR FILING DATE: 2001-12-12
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 572
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-167-994-2

Query Match 100.0%; Score 2994; DB 14; Length 572;
Best Local Similarity 100.0%; Pred. No. 2e-284;
Matches 572; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEIEIGEQPPVKCSNFFANHWKGLVFLVPLLCPLVMLLNEGAEFRCMYLLLVMAIFW 60
DB 1 MEIEIGEQPPVKCSNFFANHWKGLVFLVPLLCPLVMLLNEGAEFRCMYLLLVMAIFW 60
QY 61 VTEALPLYVTSMPIVAFPIGIMSSDQTCRLYFKDTLVNFMGGMVVALAVEYCNLHKRL 120
DB 61 VTEALPLYVTSMPIVAFPIGIMSSDQTCRLYFKDTLVNFMGGMVVALAVEYCNLHKRL 120
QY 121 ALRVIOVGCSPPRLHFLGIMVTMFLSMWISNAACTAMMCPPIQAVLEELQAGVCCKINH 180
DB 121 ALRVIOVGCSPPRLHFLGIMVTMFLSMWISNAACTAMMCPPIQAVLEELQAGVCCKINH 180
QY 181 EPQYQIVGGNKKNEDEPPYPTKITLCYILGIAYASSLGGCGTIIIGTATNLTTFKGIYEAR 240
DB 181 EPQYQIVGGNKKNEDEPPYPTKITLCYILGIAYASSLGGCGTIIIGTATNLTTFKGIYEAR 240
QY 241 FKNSTEQMDPFTFMFYSVPMSLVYLLTFVFLQWHFMGLWRPKSKEAQEVQREGADVA 300
DB 241 FKNSTEQMDPFTFMFYSVPMSLVYLLTFVFLQWHFMGLWRPKSKEAQEVQREGADVA 300
QY 301 KVIDORYKDLGPMSTHIEIQVMILFIWMVMYTRKPGIFLGWADLLNSKDINSMTPIF 360
DB 301 KVIDORYKDLGPMSTHIEIQVMILFIWMVMYTRKPGIFLGWADLLNSKDINSMTPIF 360
QY 361 VVVMCFMPLPANTAFRLYCTRRGGPVTGPTPSLITWKFIQTKVPWGLVFLGGGFALAE 420
DB 361 VVVMCFMPLPANTAFRLYCTRRGGPVTGPTPSLITWKFIQTKVPWGLVFLGGGFALAE 420
QY 421 SKQSGMAKLIGNALIGLKVLPSNVLVILVAVFLTAFSSNVAIANIIPVLAESLAI 480

RESULT 2

US-10-017-479-2
; Sequence 2, Application US/10017479
; Publication No. US20030104399A1
; GENERAL INFORMATION:
; APPLICANT: University of Connecticut
; APPLICANT: Helfand, Stephan L
; APPLICANT: Reenan, Robert A
; APPLICANT: Rogina, Blanka
; TITLE OF INVENTION: Polynucleotides Encoding Cellular Transporters and Methods of Use
; FILE REFERENCE: UCT-0020
; CURRENT APPLICATION NUMBER: US/10/017,479
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/255,013
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 572
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-017-479-2

Query Match 100.0%; Score 2994; DB 14; Length 572;
Best Local Similarity 100.0%; Pred. No. 2e-284;
Matches 572; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEIEIGEQPPVKCSNFFANHWKGLVFLVPLLCPLVMLLNEGAEFRCMYLLLVMAIFW 60
DB 1 MEIEIGEQPPVKCSNFFANHWKGLVFLVPLLCPLVMLLNEGAEFRCMYLLLVMAIFW 60
QY 61 VTEALPLYVTSMPIVAFPIGIMSSDQTCRLYFKDTLVNFMGGMVVALAVEYCNLHKRL 120
DB 61 VTEALPLYVTSMPIVAFPIGIMSSDQTCRLYFKDTLVNFMGGMVVALAVEYCNLHKRL 120
QY 121 ALRVIOVGCSPPRLHFLGIMVTMFLSMWISNAACTAMMCPPIQAVLEELQAGVCCKINH 180
DB 121 ALRVIOVGCSPPRLHFLGIMVTMFLSMWISNAACTAMMCPPIQAVLEELQAGVCCKINH 180
QY 181 EPQYQIVGGNKKNEDEPPYPTKITLCYILGIAYASSLGGCGTIIIGTATNLTTFKGIYEAR 240
DB 181 EPQYQIVGGNKKNEDEPPYPTKITLCYILGIAYASSLGGCGTIIIGTATNLTTFKGIYEAR 240
QY 241 FKNSTEQMDPFTFMFYSVPMSLVYLLTFVFLQWHFMGLWRPKSKEAQEVQREGADVA 300
DB 241 FKNSTEQMDPFTFMFYSVPMSLVYLLTFVFLQWHFMGLWRPKSKEAQEVQREGADVA 300
QY 301 KVIDORYKDLGPMSTHIEIQVMILFIWMVMYTRKPGIFLGWADLLNSKDINSMTPIF 360
DB 301 KVIDORYKDLGPMSTHIEIQVMILFIWMVMYTRKPGIFLGWADLLNSKDINSMTPIF 360
QY 361 VVVMCFMPLPANTAFRLYCTRRGGPVTGPTPSLITWKFIQTKVPWGLVFLGGGFALAE 420
DB 361 VVVMCFMPLPANTAFRLYCTRRGGPVTGPTPSLITWKFIQTKVPWGLVFLGGGFALAE 420
QY 421 SKQSGMAKLIGNALIGLKVLPSNVLVILVAVFLTAFSSNVAIANIIPVLAESLAI 480

421	SKQSGMAXLIGNALIGLGLKVLNPSVLLVVILVAVFLTAFSSNVAVIANIIPVLAEMSLAI	480
481	EIHPLYLLPLAGLACSMAPHLPVSTPPNALVAGYANIRTKMAIAGIGPTITITITLFPV	540
481	EIHPLYLLPLAGLACSMAPHLPVSTPPNALVAGYANIRTKMAIAGIGPTITITITLFPV	540
541	QQTGLVVYPNLNSFPPEWAAQIYAAAALGNKTH	572
541	QQTGLVVYPNLNSFPPEWAAQIYAAAALGNKTH	572

```

RESULT 3
US-10-718-359-2.
; Sequence 2, Application US/10718359
; Publication No. US20050095240A1
; GENERAL INFORMATION:
; APPLICANT: MEDICAL COLLEGE OF GEORGIA RESEARCH INSTITUTE, INC.
; TITLE OF INVENTION: NACT AS A TARGET FOR LIFESPAN EXPANSION AND WEIGHT REDUCTION
; FILE REFERENCE: 275 0008 0101
; CURRENT APPLICATION NUMBER: US/10/718,359
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/428,469
; PRIOR FILING DATE: 2002-11-22
; PRIOR APPLICATION NUMBER: 60/459,441
; PRIOR FILING DATE: 2003-04-01
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 572
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-718-359-2

```

```

QY 541 CQTGGLVVYPNLNSFPEWAIYAAAAALGNKTH 572
      |||||||
Db 541 CQTGGLVVYPNLNSFPEWAIYAAAAALGNKTH 572
      |||||||

RESULT 4
US-10-167-994-10
; Sequence 10, Application US/10167994
; Publication No. US20030082647A1
; GENERAL INFORMATION:
; APPLICANT: Reenan, Robert A.
; APPLICANT: Rogina, Blanka
; APPLICANT: Helfand, Stephen L.
; TITLE OF INVENTION: TRANSPORTER PROTEIN
; FILE REFERENCE: 13407-013001
; CURRENT APPLICATION NUMBER: US/10/167,994
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/255,013
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 10/017,479
; PRIOR FILING DATE: 2001-12-12
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-167-994-10

```

Query Match	100.0%	Score 2994	DB 17	Length 572
Best Local Similarity	100.0%	Mismatches 0	Indels 0	Gaps 0
Matches 572	Conservative 0	Partial 284		
Qy	1	MEIEIGQOPQPVKCSNFFFAHNMGLVFLVPLLCPLPVMLNNEGAFRCMYLLVMAFW	60	
Db	1	MEIEIGQOPQPVKCSNFFFAHNMGLVFLVPLLCPLPVMLNNEGAFRCMYLLVMAFW	60	
Qy	61	VTEALPLYVTSMIPIVAPPIINGIMSSDOTCRLYFKDTLVMFMGGIMVALAVEYCNLHKRL	120	
Db	61	VTEALPLYVTSMIPIVAPPIINGIMSSDOTCRLYFKDTLVMFMGGIMVALAVEYCNLHKRL	120	
Qy	121	ALRVIQIVGCSPRRLHFGILVMTWFLSMWISNAACTAMMCPIIQAVLBELOAQGVCKINH	180	
Db	121	ALRVIQIVGCSPRRLHFGILVMTWFLSMWISNAACTAMMCPIIQAVLBELOAQGVCKINH	180	
Qy	181	EPQYQIVGGNKKNEDEPPYPTKTLTCYYLGIAYASSIGGGCTIIGTATNTLTFKGIYEAR	240	
Db	181	EPQYQIVGGNKKNEDEPPYPTKTLTCYYLGIAYASSIGGGCTIIGTATNTLTFKGIYEAR	240	
Qy	241	FKNSTEQMDPPTFMFYSPVPSMLVYTLTFVLQWHFGLWRPKSKSEAQEVQREGADV	300	
Db	241	FKNSTEQMDPPTFMFYSPVPSMLVYTLTFVLQWHFGLWRPKSKSEAQEVQREGADV	300	
Qy	301	KKVIDQRYKDLGPMSSIHIEIQWILFIEMVMYFTRKPGIFIGWADLLNSKDIRNSMPTIF	360	
Db	301	KKVIDQRYKDLGPMSSIHIEIQWILFIEMVMYFTRKPGIFIGWADLLNSKDIRNSMPTIF	360	
Qy	361	VVMCFMLPANYAFRLRYCTRGGVPVPTPSLITWKIEIQTKVPMLGVFLIGGGFALAE	420	
Db	361	VVMCFMLPANYAFRLRYCTRGGVPVPTPSLITWKIEIQTKVPMLGVFLIGGGFALAE	420	
Qy	421	SKOSGMAKLIGNALIGLKVLPNSVLLLVILVAVFLTAFSSNVAIANIIIPVLAEMSLAI	480	
Db	421	SKOSGMAKLIGNALIGLKVLPNSVLLLVILVAVFLTAFSSNVAIANIIIPVLAEMSLAI	480	
Qy	481	EIHPLYLILPAGLACSMAFHLPVSTPPNALVAGYANIRTKDMAIAGIPTIITITLFFV	540	
Db	481	EIHPLYLILPAGLACSMAFHLPVSTPPNALVAGYANIRTKDMAIAGIPTIITITLFFV	540	

RESULT 5

US-10-173-519-5
; Sequence 5, Application US/10173519
; Publication No. US20020193582A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Millennium Pharmaceuticals, Inc.
; TITLE OF INVENTION: 69624, A Human Transporter Family Member
; TITLE OF INVENTION: and Uses Therefor
; FILE REFERENCE: MPI01-098PIRNM
; CURRENT APPLICATION NUMBER: US/10/173,519
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 60/298,970
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 587
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-173-519-5

Query Match 34.8%; Score 1042.5; DB 13; Length 587;
Best Local Similarity 38.1%; Pred. No. 5.9e-93;
Matches 219; Conservative 129; Mismatches 190; Indels 37; Gaps 8;

QY 26 LVVFLVLLCLPVMNLNEGAEFRCMYLLVMAIFWVTEALPLYVTSMIPVAPINGIMS 85
DB 15 LIVLCLPIFLPLLIIVQTKAEYCAYSIILMALLWCTEALPLAVTALFPIVLPFLMGIND 74

QY 86 SDQTCRLYFKDTLVMFMGGIMVALAVCYCNLHKRLALRVIOIGVCSRRRLHFGIMVTMF 145
DB 75 ASEVCIEYFKDTNLFVGGIMVALAVEHWNHKLRIALQVLLIIGVRPALLLGFMLVTAF 134

QY 146 LSMWISNACTAMMCPIIQAVLELQA--QGVCKINHEPOYQIV-----GGNK 191
DB 135 LSMWISNTATTAMVPIGHAVLEQLQSKKQVEGNNNPTFELQEECPQKEVTKLDNGQP 194

QY 192 KNEDEP-----PYPTKITLCYVLGIAYASSLGCGCTIIGTATNLTFRGIYEAREFK 242
DB 195 VSAPSEPRTKTQEHHRFSQGLSLC-----ICYSASIGGIATLTGTPNLVLOGVNSLFP 250

QY 243 NSTEQMDPFTFMYSVPSMLVYLLTFVFLQWHFMGLWRPKSEAOEVORGREGADVAKK 302
DB 251 QNGNVNFAFWGFGAFTMIILLLLAWLQVLFLGVNFRKNFGFGEGERKQA--AFQ 308

QY 303 VIDQRYKDLGPMSTHEIQVMILFIMVYFTRKPGIFLGWADLLNSKDIRNSMPT---- 358
DB 309 VIKTQRYLLGPMSPAEKTVTLVFLVLLVWFTREPFGFPFGMGDTVFANEKGQSMASDGTV 368

QY 359 -IFVVMCFMLPANYAFRLYCTRGGVPVPTPSLTITWKFTIQTQKVPWGLVFLGGGFAL 417
DB 369 AIFISLVMFIIPSKIPGLMQDPKPKGLKA--PPAILTWKTVDNRMPNIVILLGGGFAL 426

QY 418 AEGSQSGMAKLGNALIGLKVLPNSVLLVAVELTAFSSNVAIANIIPVLAEMS 477
DB 427 AKGSEQSLSEWLGDKLTPLQHPSPSATAVILCLLTAIFTECTSNVATTLFLPILASMA 486

QY 478 LAIEIHPYLIILPAGLACSMAFHLPVSTPPNALVAGYANIRTKDMAIAGIGPTIITL 537
DB 487 QAICLHPYVMLPCTLAASLAFMLPVATPPNAIVFSGGLKVDMARAGFLNIIIGVLAI 546

QY 538 FVFCQWGLVVPYVNLNSFPPEWAIYAAAALGNKTH 572
DB 547 TLSINSWSIPIF-KLDTFFPSWAHSNTSQCLNPSN 580

RESULT 6

US-10-167-994-12
; Sequence 12, Application US/10167994
; Publication No. US20030082647A1

; GENERAL INFORMATION:

; APPLICANT: Reenan, Robert A.
; APPLICANT: Rogina, Blanka
; APPLICANT: Helfand, Stephen L.
; TITLE OF INVENTION: TRANSPORTER PROTEIN
; FILE REFERENCE: 13407-013001
; CURRENT APPLICATION NUMBER: US/10/167,994
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/255,013
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 10/017,479
; PRIOR FILING DATE: 2001-12-12
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 587
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-167-994-12

Query Match 34.8%; Score 1042.5; DB 14; Length 587;
Best Local Similarity 38.1%; Pred. No. 5.9e-93;
Matches 219; Conservative 129; Mismatches 190; Indels 37; Gaps 8;

QY 26 LVVFLVLLCLPVMNLNEGAEFRCMYLLVMAIFWVTEALPLYVTSMIPVAPINGIMS 85
DB 15 LIVLCLPIFLPLLIIVQTKAEYCAYSIILMALLWCTEALPLAVTALFPIVLPFLMGIND 74

QY 86 SDQTCRLYFKDTLVMFMGGIMVALAVCYCNLHKRLALRVIOIGVCSRRRLHFGIMVTMF 145
DB 75 ASEVCIEYFKDTNLFVGGIMVALAVEHWNHKLRIALQVLLIIGVRPALLLGFMLVTAF 134

QY 146 LSMWISNACTAMMCPIIQAVLELQA--QGVCKINHEPOYQIV-----GGNK 191
DB 135 LSMWISNTATTAMVPIGHAVLEQLQSKKQVEGNNNPTFELQEECPQKEVTKLDNGQP 194

QY 192 KNEDEP-----PYPTKITLCYVLGIAYASSLGCGCTIIGTATNLTFRGIYEAREFK 242
DB 195 VSAPSEPRTKTQEHHRFSQGLSLC-----ICYSASIGGIATLTGTPNLVLOGVNSLFP 250

QY 243 NSTEQMDPFTFMYSVPSMLVYLLTFVFLQWHFMGLWRPKSEAOEVORGREGADVAKK 302
DB 251 QNGNVNFAFWGFGAFTMIILLLLAWLQVLFLGVNFRKNFGFGEGERKQA--AFQ 308

QY 303 VIDQRYKDLGPMSTHEIQVMILFIMVYFTRKPGIFLGWADLLNSKDIRNSMPT---- 358
DB 309 VIKTQRYLLGPMSPAEKTVTLVFLVLLVWFTREPFGFPFGMGDTVFANEKGQSMASDGTV 368

QY 359 -IFVVMCFMLPANYAFRLYCTRGGVPVPTPSLTITWKFTIQTQKVPWGLVFLGGGFAL 417
DB 369 AIFISLVMFIIPSKIPGLMQDPKPKGLKA--PPAILTWKTVDNRMPNIVILLGGGFAL 426

QY 418 AEGSQSGMAKLGNALIGLKVLPNSVLLVAVELTAFSSNVAIANIIPVLAEMS 477
DB 427 AKGSEQSLSEWLGDKLTPLQHPSPSATAVILCLLTAIFTECTSNVATTLFLPILASMA 486

QY 478 LAIEIHPYLIILPAGLACSMAFHLPVSTPPNALVAGYANIRTKDMAIAGIGPTIITL 537
DB 487 QAICLHPYVMLPCTLAASLAFMLPVATPPNAIVFSGGLKVDMARAGFLNIIIGVLAI 546

QY 538 FVFCQWGLVVPYVNLNSFPPEWAIYAAAALGNKTH 572
DB 547 TLSINSWSIPIF-KLDTFFPSWAHSNTSQCLNPSN 580

RESULT 7

US-10-718-359-13
; Sequence 13, Application US/10718359
; Publication No. US20050095240A1
; GENERAL INFORMATION:
; APPLICANT: MEDICAL COLLEGE OF GEORGIA RESEARCH INSTITUTE, INC.
; TITLE OF INVENTION: NACT AS A TARGET FOR LIFESPAN EXPANSION AND WEIGHT REDUCTION
; FILE REFERENCE: 275.0008 0101

; CURRENT APPLICATION NUMBER: US/10/718,359
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/428,469
; PRIOR FILING DATE: 2002-11-22
; PRIOR APPLICATION NUMBER: 60/459,441
; PRIOR FILING DATE: 2003-04-01
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 13
; TYPE: PRT
; ORGANISM: rat NaDC1
US-10-718-359-13

Query Match 34.8%; Score 1042.5; DB 17; Length 587;
Best Local Similarity 38.1%; Pred. No. 5,9e-93;
Matches 219; Conservative 129; Mismatches 190; Indels 37; Gaps 8;
QY 26 LVVFLVPLCLPVMMLNEGAEPRCMLLVLLVMAIFWVTEALPLYVTSMIPVIAFPINGIMS 85
DB 15 LIVLCLPIPLPLPIVOTKEAYCAYSIILMALLWCTEALPLAVTALFVILFPLMGIMD 74
QY 86 SDOTCRLYFKDTLVMFMGGMVALAVEYCNLHKRLALRVIOIVGCSPRHLHGLIWTMP 145
DB 75 ABEVCIEYFKDTNLIIFVGLMVAIAVEYCNLHKRLALQVLLIIGVRPALLLLGFMLVTAF 134
QY 146 LSMWISNAACTAMMCPITQAVLEELQA--QGVCKINHEPQYQIV-----GNGK 191
DB 135 LSMWISNTATTAMVPIGHAVLEQLQSGKDVGGNNPTFELQECQKEVTKLDNQOP 194
QY 192 KNEDEP-----PVTKITLVCYLGIAVASSLGGCGTIIIGTATNLTFRKGIYEAFPK 242
DB 195 VSAPSEPRQTQEHHRFSQGLSLC---ICYSASIGGIATLTGTTPNLVQGVNSLFP 250
QY 243 NSTEQMDPTFMFYSVPSMLVYTLITFVFLQWHFGLWRPKSKAEQVORGEGADVAKK 302
DB 251 QNGNVNFASFPGFAPPTMIILLLAWLQVLFGLGVNFRKNFGFGEGERKQA--AFQ 308
QY 303 VIDORYKDLGPMISHEIOVMILFIWMVYFTRKPGIFLGMADLNSKDINSMPT---- 358
DB 309 VIKTYRLLGPNSPAETVTVLVLLVNLVWTFREGFPGFGCDTVFANEKGOSMADGTV 368
QY 359 -IFVVMCFMPLPANYAFILRYCTRGGPVPTGTPSLITWKFTQTKVPKGLVFLGGG 417
DB 369 AIFISLVNFIIPSKIPGLMQDKPKGLKA--PPAILTWKTWNQWNNVILVGGG 426
QY 418 AEGSQSGMAKLIAGNALGLKVLNLSVLLVILVAVFLTAFSSNVATANIIPVLAEMS 477
DB 427 AKGSEQSGLEWLGKLTPLQHIPPSATAVILCLLIAIFTECTSNVATTLFLPLASMA 486
QY 478 LAIEIHPYLIPLAGLACSMAPHLVPSTPPNALVAGYANIRTKMAIAGIGPTIITL 537
DB 487 QAICLHPYLVLPCTLASLAPLVPATPPNAIVFSFGCLKVSDMARAGFLNIIGVLAI 546
QY 538 FVFCOTGLVVPNLNSPPEWAAIAALGNKTH 572
DB 547 TLSINSWSIPIF-KLDTFSPWAHSNTSQCLNPSN 580

RESULT 8
US-10-718-359-12
; Sequence 12, Application US/10718359
; Publication No. US20050095240A1
; GENERAL INFORMATION:
; APPLICANT: MEDICAL COLLEGE OF GEORGIA RESEARCH INSTITUTE, INC.
; TITLE OF INVENTION: NACT AS A TARGET FOR LIFESPAN EXPANSION AND WEIGHT REDUCTION
; FILE REFERENCE: 275,008 0101
; CURRENT APPLICATION NUMBER: US/10/718,359
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/428,469
; PRIOR FILING DATE: 2002-11-22
; PRIOR APPLICATION NUMBER: 60/459,441
; PRIOR FILING DATE: 2003-04-01

; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12
; LENGTH: 581
; TYPE: PRT
; ORGANISM: zebrafish NaCT
US-10-718-359-12
Query Match 34.7%; Score 1039; DB 17; Length 581;
Best Local Similarity 38.0%; Pred. No. 1.3e-92;
Matches 221; Conservative 117; Mismatches 185; Indels 58; Gaps 10;
QY 23 WK---GLVFLVPLCLPVMMLNEGAEPRCMLLVLLVMAIFWVTEALPLYVTSMIPVIAFP 79
DB 10 WKMKNTLILFCTPFLPLPLPLVIGSKAGACAYVVMVMAVYVCTEVLPLAVTALLPAVLFP 69
QY 80 IMGIMSSDOTCRLYFKDTLVMFMGGMVALAVEYCNLHKRLALRVIOIVGCSPRHLHGL 139
DB 70 LFRIMESQDVCMQVYKDKTNMLFGLGLMVAIAVEYCNLHKRLALRVLLLVGVRPALLMLGF 129
QY 140 IMVTWFLSMWISNAACTAMMCPITQAVLEELQAQGVCKINHEPQYQIVGCKNKNEDP 199
DB 130 MGVTAFLSMWISNTATTAMVPIVQAVLEQLN-----NTAQEQSSIFETEEKSTEKQPE 184
QY 200 YPTK-----ITLCYYLGIAYASSLGGCGTIIIGT 227
DB 185 SPGEKVVNLGDNFMSBDDPEHSREABERLKMKGTLTC----VCYAASIGGTATLTGT 240
QY 228 ATNLTFRKGIYEAFKPNSTEQMDPTFMFYSVPSMLVYTLITFVFLQWHFGLWRPKSKEA 287
DB 241 GPNLVLMGMSQLPDPDIIINFASWFGFAPFNMIIMLTLAWLMLQIVFLGINKFKTWGC 300
QY 288 QEVORGEGADVAKKIDORYKDLGPMISHEIOVMILFIWMVYFTRKPGIFLGMADL 347
DB 301 GTVTEKIEA--AYNVKEHRSGLPMTFGBLSVALFILVVLWTFDRDGFVDGWAIRF 358
QY 348 NSKD---IRNSMPTIFVVMCFMPLPANYAFILRYC---TRRGSPV---TGTPPSLITWK 398
DB 359 FNADKEFVTDATVAVFVAALLFVFPKPP--RLCFWRTESEDTVFOQESGPTPALLTWKV 416
QY 399 IOTKVPKGLVFLGGGALAEQSGSKMAKLIAGNALGLKVLNLSVLLVILVAVFLTA 458
DB 417 TQKMPMSIILLGGGALAKGSEISGLSKWLGQMSPLQSIPPMAIAIVICLMIATFTE 476
QY 459 FSSNVATANIIPVLAEMSLAIEIHPYLIPLAGLACSMAPHLVPSTPPNALVAGYANIR 518
DB 477 CTSNVATATFLPLASMSQSIGNVPLVYVWVPCVTLASFAFAPLVPATPPNAIVFSYGLK 536
QY 519 TKDMAIAGIGPTIITLITFVFCOTGLVVPNLNSPPEWA 559
DB 537 VSDMAKTGIVNMIIGILSITLAINSWGRAIF-SLDTFPSWA 576

RESULT 9
US-10-167-994-3
; Sequence 3, Application US/10167994
; Publication No. US20030082647A1
; GENERAL INFORMATION:
; APPLICANT: Reenan, Robert A.
; APPLICANT: Rogina, Blanka
; APPLICANT: Helfand, Stephen L.
; TITLE OF INVENTION: TRANSPORTER PROTEIN
; FILE REFERENCE: 13407-013001
; CURRENT APPLICATION NUMBER: US/10/167,994
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/255,013
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 10/017,479
; PRIOR FILING DATE: 2001-12-12
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 592

```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-167-994-3

Query Match      34.3%; Score 1026; DB 14; Length 592;
Best Local Similarity 37.2%; Pred. No. 2.5e-91;
Matches 219; Conservative 124; Mismatches 189; Indels 56; Gaps 10;

QY 26 LVVFLVPLLCPLPVMMLNEGAFRCMYLLLVMAIFWVTEALPLVYTSIMPIVAFPIGIMS 85
DB 15 LIVFFVPIILLPLPLVPSKEAYCAVAILMALFWCTEALPLAVTALFPLILFPMGIVD 74

QY 86 SDQTCRLYKDTLVNFMGGIMVALAVEYCNLHKRLALRVIOIGCSPPRLHFLGLIMVTF 145
DB 75 ASEVAVEYLKDSNLLFFGGLLVAVIAVEHNLHKRLALRVLLIVGVRPAPLILGMLVTAF 134

QY 146 LSMWISNAACTAMMCPITQAVLEELQAQGVCKINHEPOYQIVGKNKKNDEDEPPYTK-- 203
DB 135 LSMWISNTATSAMVPIAHAVLDQLHSSQASSNVEE-----GSNNPTFELQEPSQKEV 188

QY 204 -----ITLCYYLGIAYASSLGCGCTIIGTATNLTFTKG 235
DB 189 TKLDNGQALPVTSSASSEGRAHLSQKHLHLTQCMSLCVCYSASIGGIATLTGTAPNLVLQ 248

QY 236 IYERFKNSTEQMDPTFMFYSVPSMLVYLLTFVFLQHFHMLG-WRPKSKAEQVQRGR 294
DB 249 QINSIFPQNGVNVFASFSEFAPFTWVILLLLAWLQILFLGFNFRKNFGIGEMQ--- 305

QY 295 EGADVAKKVIDQRYKDLGPMSTHIEIQVMILFIFMVVMYFTRKPGIFLGWADLL--NSKD- 351
DB 306 EQQAAAYCVIQTHERLLGPMWTFEAKAISILFVILVLLWTFREPGFPLGNGNLAFPNKAGE 365

QY 352 --IRNSMPTIFVVMCFMPLPANYAFLYRYCTRRGGPV--PTGPTPSLITWKFIQTKVPWGL 407
DB 366 SMVSDGTVAIFIGIIMFIIPSKFPGLTQDPENPGKLKAPLG---LLDWKTVNQKMPWNI 421

QY 408 VFLLGGGFALAGSKQSGMAKLGNALIGLVLPNSVLLVLLVAVLFAVFLTAFFSSNVAIAN 467
DB 422 VLLGGYALAGSKRSGLSEWLNKLTPLQSVPAIPAIAIILSLVATFTECTSNVATT 481

QY 468 IIPVLAEMSLAIEIHPYLIILPAGLACSMAFHLPVSTPPNALVAGYANIRTKDMAIAGI 527
DB 482 IFLPILASMAQAICLHPLVYMLPCTLATSLAFMLPVATPPNAIVFSPGDLKVLDMARAGF 541

QY 528 GPTITITIFVFCQWGLVVPYVNLNGPPEWQAIIYAAA---ALGNKT 571
DB 542 LLNIIGVLIILAINSWGIPLF-SLHSPFSPAQSNTTAQCLPSLANTT 588

RESULT 11
US-09-729-094-4
; Sequence 4, Application US/09729094
; Patent No. US20020019028A1
; GENERAL INFORMATION:
; APPLICANT: CHATURVEDI, Kabir et al
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000662
; CURRENT APPLICATION NUMBER: US/09/729,094
; CURRENT FILING DATE: 2000-12-05
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 619
; TYPE: PRT
; ORGANISM: Xenopus laevis
US-09-729-094-4

Query Match      33.8%; Score 1010.5; DB 9; Length 619;
Best Local Similarity 35.3%; Pred. No. 8.8e-90;
Matches 218; Conservative 120; Mismatches 192; Indels 87; Gaps 7;

QY 17 NFFANHWKGLVYVFLPLCLPVMMLNEGAFRCMYLLLVMAIFWVTEALPLVYTSIMPIV 76
DB 13 NYF-----IIFLPLFLPLPLVVPYVNLNGPPEWQAIIYAAA---ALGNKT 65

```

```
QY 77 APPINGMSDQTCRLYFKDTLVMFMGGIMVALAVEYCNLHKRLALRVQIVGCSPPRLH 136
DB 66 LPPMMGIMDSQTCRLYFKDTLVMFMGGIMVALAVEYCNLHKRLALRVQIVGCSPPRLH 125
QY 137 FGLIMVTFMFLSMWISNAACATAMCPITIOAVLEEL----- 170
DB 126 LGFMVVTAFLSNMWISNATATTAMPIAOAVMEQLHSSEKQVDERVEGNSNTOKNVNGHEN 185
QY 171 ----- 240
DB 186 DMYESVMPGKMAIDNTYATENEGFEIQEKSTKDPSPKQEKQSIGPIVIEPE----- 240
QY 189 GNKKNEDEPPPTKTLTCYLYGIAYASSLGGCGTIICTATNLTPKGIYEAFPKNSTEQM 248
DB 241 DEKQTEEKQEKHLKICKGMSLCVCYSASIGGIATLTGTTNPLVMKGQMDLFPENNII 300
QY 249 DPPTFMFVSVMVYLLTTFVFLQWFMGLWRPKSKEAQEVQORREGADVAKKVIDQRY 308
DB 301 NFASWFGFAFTMLVLLALSWLQFIYLVGNFKNFCGCGNABQEKKEKRAFRVISGEH 360
QY 309 KDLGPMSTHEIQVMILFIFMVVMYFTRKPGIFLGMADLLNSKD----IRNSMPTIFVVM 364
DB 361 KGLSMTFAEISVLVLFILLVLLMFTREPMPGPNATISFNKGGKEMVTDATVAIFVSLM 420
QY 365 CFMLPANYAFRLY-CTRRGGPVP-TGPTPSLITWKFIQTKVPWGLVFLGGGFALAEKSK 422
DB 421 MFFPSELPSFKYQDTPKPGMKPLRVPPALLDMKTVNEKMPWNIIVLLGGGFALAKGSE 480
QY 423 QSGMAKLIGNALIGLVKLVNSVLLVILVAVFLTAFSSNVAIANIIPVLAEMSLAIEI 482
DB 481 EBSGLMWGEKLTPLQSIPTPAIAIALICLLVATFTECTSNVATTTFLFPLASMAKAIQL 540
QY 483 HPLYLILPAGLACSMFHLPVSTPPNALVAGYANIRTKDMAIAGIPTIITITLFFVFCQ 542
DB 541 NPLYIMLPCTLSASLAFMLPVATPPNATFSAFSGQLKVIDMAKAGLLNLGLVLTITLAIN 600
QY 543 TWGLVYVYNLNSFPPEWA 559
DB 601 SWGFYMF-NLGTFFPSWA 616

RESULT 12
US-10-435-631-4
; Sequence 4, Application US/10435631
; Publication No. US20030186381A1
; GENERAL INFORMATION:
; APPLICANT: CHATURVEDI, Kabir et al
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000662CON
; CURRENT APPLICATION NUMBER: US/10/435,631
; CURRENT FILING DATE: 2003-05-12
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 619
; TYPE: PRT
; ORGANISM: Xenopus laevis
US-10-435-631-4

Query Match 33.8%; Score 1010.5; DB 14; Length 619;
Best Local Similarity 35.3%; Pred. No. 8,8e-90;
Matches 218; Conservative 120; Mismatches 192; Indels 87; Gaps 7;
```

```
QY 17 NFFANHWKGLVFLVPLCLPMLNEGABRCMYLLLVMAIFWVTEALPLVVTSMIPIV 76
DB 13 NYF-----IIFLVPFLPLPLVVPVTKASCGFVIVVMAIFWVTEALPLVLTALFPVL 65
QY 77 APPINGMSDQTCRLYFKDTLVMFMGGIMVALAVEYCNLHKRLALRVQIVGCSPPRLH 136
DB 66 LPPMMGIMDSQTCRLYFKDTLVMFMGGIMVALAVEYCNLHKRLALRVQIVGCSPPRLH 125
```

```
QY 137 FGLIMVTFMFLSMWISNAACATAMCPITIOAVLEEL----- 170
DB 126 LGFMVVTAFLSNMWISNATATTAMPIAOAVMEQLHSSEKQVDERVEGNSNTOKNVNGHEN 185
QY 171 ----- 240
DB 186 DMYESVMPGKMAIDNTYATENEGFEIQEKSTKDPSPKQEKQSIGPIVIEPE----- 240
QY 189 GNKKNEDEPPPTKTLTCYLYGIAYASSLGGCGTIICTATNLTPKGIYEAFPKNSTEQM 248
DB 241 DEKQTEEKQEKHLKICKGMSLCVCYSASIGGIATLTGTTNPLVMKGQMDLFPENNII 300
QY 249 DPPTFMFVSVMVYLLTTFVFLQWFMGLWRPKSKEAQEVQORREGADVAKKVIDQRY 308
DB 301 NFASWFGFAFTMLVLLALSWLQFIYLVGNFKNFCGCGNABQEKKEKRAFRVISGEH 360
QY 309 KDLGPMSTHEIQVMILFIFMVVMYFTRKPGIFLGMADLLNSKD----IRNSMPTIFVVM 364
DB 361 KGLSMTFAEISVLVLFILLVLLMFTREPMPGPNATISFNKGGKEMVTDATVAIFVSLM 420
QY 365 CFMLPANYAFRLY-CTRRGGPVP-TGPTPSLITWKFIQTKVPWGLVFLGGGFALAEKSK 422
DB 421 MFFPSELPSFKYQDTPKPGMKPLRVPPALLDMKTVNEKMPWNIIVLLGGGFALAKGSE 480
QY 423 QSGMAKLIGNALIGLVKLVNSVLLVILVAVFLTAFSSNVAIANIIPVLAEMSLAIEI 482
DB 481 EBSGLMWGEKLTPLQSIPTPAIAIALICLLVATFTECTSNVATTTFLFPLASMAKAIQL 540
QY 483 HPLYLILPAGLACSMFHLPVSTPPNALVAGYANIRTKDMAIAGIPTIITITLFFVFCQ 542
DB 541 NPLYIMLPCTLSASLAFMLPVATPPNATFSAFSGQLKVIDMAKAGLLNLGLVLTITLAIN 600
QY 543 TWGLVYVYNLNSFPPEWA 559
DB 601 SWGFYMF-NLGTFFPSWA 616

RESULT 13
US-10-173-519-2
; Sequence 2, Application US/10173519
; Publication No. US20020193582A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
; TITLE OF INVENTION: 69624, A Human Transporter Family Member
; TITLE OF INVENTION: and Uses Therefor
; FILE REFERENCE: MP101-098P1RNM
; CURRENT APPLICATION NUMBER: US/10/173,519
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 60/298,970
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 568
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-173-519-2

Query Match 33.5%; Score 1004; DB 13; Length 568;
Best Local Similarity 37.2%; Pred. No. 3,4e-89;
Matches 213; Conservative 121; Mismatches 193; Indels 46; Gaps 11;
```

```
QY 19 FANHWKGLV-FLVPLCLPMLNEGABRCMYLLLVMAIFWVTEALPLVVTSMIPIVA 77
DB 7 YVSKFKSFILFVTPVPLLLPLVILMPAKFVRCVAVIILMAIYCTEVIPLVATSLMPVLL 66
QY 78 FPIINGMSDQTCRLYFKDTLVMFMGGIMVALAVEYCNLHKRLALRVQIVGCSPPRLH 137
DB 67 FPLFQILDSRQVCVQYMKDITNMLFLGLIIVAVAVERNLHKRIALRTLLVWGAKEPARML 126
QY 138 GLIMVTFMFLSMWISNAACATAMCPITIOAVLEELQAGVC-----KINHEPOVQI 186
```

Db 127 GFMGTALLSMWISNTATTAMVPIVEAILQOMBEATSAATEAGLELVKDGKAKELPGSQV 186
QY 187 VGGNKKNNEDPPPTKTIILCYILGIAYASSLGGGIIIGTATNLTKGIYEAPFKNSTE 246
Db 187 IFEGTLLGQOQDQERKRLCKAMTLCICVAAISIGGTATLTGTGPNVLLGQMNELFPDSKD 246
QY 247 QMDPFTFMFYSVPSMLVYTLTLTFVFLQWHFM-----GLWRPKSKEAQEVQREG 296
Db 247 LVNFASFASFAPFNMLVLLFAWLWLFQVYMFNFKXSGCGLESKNEKA----- 297
QY 297 ADVAKKVIDQRYKOLGPMISHEIQVMILFIMVVMYFTRKPGIFLGWADLL-----NSKDI 352
Db 298 ---ALKVLOEYRKLGLPLSFAEINVLICFFLLVILWFSRDPGFMGLTVAWVEGETKYV 354
QY 353 RNSMPTIFVVMCEMLPA---NYAFRLYCTRRGGPVPTGPTPSLITWKFTQTKVPMGLVF 409
Db 355 SDATVAIFVATLLFIVPSQKPFNF-RSQTEBEERTPFYP-PPLDDWKVTOEKVPWGI 412
QY 410 LLGGGFALAEKSGKQSMKLIIGNALIGLVKLPNSVLLLV-ILVAVFLTAFSSNVAIANI 468
Db 413 LLGGGFALAKGSEASGLSVWNGKQMEPLHAVPPAAITLILSLLVAVP-TECTSNVATTL 471
QY 469 IIPVLAEMSLAIEIHPLYLILPAGLACSMAPHLPVSTPPNALVAGYANIRTKDMAIAGIG 528
Db 472 FLPIFASMSRSIGLNPLIYMLPCTLSASFAPMLPVATPPNAIVFTYGHKLKVDAMVKTGVI 531
QY 529 PTIITITILFVFCOTWGLVVPNLNSPPEWAOI 561
Db 532 MNIIIGVFCVFLAVNTWGRAIF-DLDHFPDWMNV 563

RESULT 14

US-10-403-161-72
; Sequence 72, Application US/10403161
; Publication No. US2004004930A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David et al.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-573C
; CURRENT APPLICATION NUMBER: US/10/403,161
; CURRENT FILING DATE: 2003-03-31
; PRIOR APPLICATION NUMBER: 60/370349
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/384543
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/370969
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 60/403748
; PRIOR FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: 60/372019
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/374379
; PRIOR FILING DATE: 2002-04-22
; PRIOR APPLICATION NUMBER: 09/779679
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/181045
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: 10/055877
; PRIOR FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: 60/262892
; PRIOR FILING DATE: 2001-01-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: Curaseq1 version 0.1
; SEQ ID NO 72
; LENGTH: 568
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-403-161-72

Query Match 33.5%; Score 1004; DB 15; Length 568;
Best Local Similarity 37.2%; Pred. No. 3.4e-89;
Matches 213; Conservative 121; Mismatches 193; Indels 46; Gaps 11;

QY 19 FANHKKGLVV-FLVPLLCIPVMLINEGAEFRCMYLLVMAIFWVTEALPLYVTSMPIPIVA 77
Db 7 YVSKFKSFVILFTPTLLPLVLIMPAKFRVCAYVILMAIYWCTEVIPLAVTSLMEVLL 66
QY 78 FPIGMISMSDQTCRLRYFKDTLVNFMGGIMVALAVEYCNLHKRLALRVIOIVGCSPPRLHF 137
Db 67 FPLFQILDSROVCVQMKDQNMFLFGGLIVAVAVERNLHKRIALRTLILWVGAKPARLML 126
QY 138 GLIMVTMTSMWISNTAATMMCPPIIOAVLEELQAQGV-----KINHEPQYQI 186
Db 127 GFMGTALLSMWISNTATTAMVPIVEAILQOMBEATSAATEAGLELVKDGKAKELPGSQV 186
QY 187 VGGNKKNNEDPPPTKTIILCYILGIAYASSLGGGIIIGTATNLTKGIYEAPFKNSTE 246
Db 187 IFEGTLLGQOQDQERKRLCKAMTLCICVAAISIGGTATLTGTGPNVLLGQMNELFPDSKD 246
QY 247 QMDPFTFMFYSVPSMLVYTLTLTFVFLQWHFM-----GLWRPKSKEAQEVQREG 296
Db 247 LVNFASFASFAPFNMLVLLFAWLWLFQVYMFNFKXSGCGLESKNEKA----- 297
QY 297 ADVAKKVIDQRYKOLGPMISHEIQVMILFIMVVMYFTRKPGIFLGWADLL-----NSKDI 352
Db 298 ---ALKVLOEYRKLGLPLSFAEINVLICFFLLVILWFSRDPGFMGLTVAWVEGETKYV 354
QY 353 RNSMPTIFVVMCEMLPA---NYAFRLYCTRRGGPVPTGPTPSLITWKFTQTKVPMGLVF 409
Db 355 SDATVAIFVATLLFIVPSQKPFNF-RSQTEBEERTPFYP-PPLDDWKVTOEKVPWGI 412
QY 410 LLGGGFALAEKSGKQSMKLIIGNALIGLVKLPNSVLLLV-ILVAVFLTAFSSNVAIANI 468
Db 413 LLGGGFALAKGSEASGLSVWNGKQMEPLHAVPPAAITLILSLLVAVP-TECTSNVATTL 471
QY 469 IIPVLAEMSLAIEIHPLYLILPAGLACSMAPHLPVSTPPNALVAGYANIRTKDMAIAGIG 528
Db 472 FLPIFASMSRSIGLNPLIYMLPCTLSASFAPMLPVATPPNAIVFTYGHKLKVDAMVKTGVI 531
QY 529 PTIITITILFVFCOTWGLVVPNLNSPPEWAOI 561
Db 532 MNIIIGVFCVFLAVNTWGRAIF-DLDHFPDWMNV 563

RESULT 15

US-10-490-080-1
; Sequence 1, Application US/10490080
; Publication No. US20040253597A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.
; TITLE OF INVENTION: Novel Protein and its DNA
; FILE REFERENCE: P02-0109PCT
; CURRENT APPLICATION NUMBER: US/10/490,080
; CURRENT FILING DATE: 2004-03-17
; PRIOR APPLICATION NUMBER: JP 2001-281992
; PRIOR FILING DATE: 2001-09-17
; PRIOR APPLICATION NUMBER: JP 2001-306873
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: JP 2002-113279
; PRIOR FILING DATE: 2002-04-16
; NUMBER OF SEQ ID NOS: 42
; SEQ ID NO 1
; LENGTH: 568
; TYPE: PRT
; ORGANISM: Human
US-10-490-080-1

Query Match 33.5%; Score 1004; DB 16; Length 568;
Best Local Similarity 37.2%; Pred. No. 3.4e-89;
Matches 213; Conservative 121; Mismatches 193; Indels 46; Gaps 11;

QY 19 FANHKKGLVV-FLVPLLCIPVMLINEGAEFRCMYLLVMAIFWVTEALPLYVTSMPIPIVA 77
Db 7 YVSKFKSFVILFTPTLLPLVLIMPAKFRVCAYVILMAIYWCTEVIPLAVTSLMEVLL 66


```
Qy 78 FPMGIMSSDOTCRLYPKDTLWFMGIMVALAVEYCNLHKRLALRVIOIVGSPRRLHF 137
Db 67 FPLFQILDSROVCVQYMKDTNMLFLGLIIVAVAVERNLHKRIALRTLWVGAKPARLML 126
Qy 138 GLIMVTMFLSMWISNAACACTAMCPIIOAVLELELOAGVC-----KINHEPOYQI 186
Db 127 GFVGVTALLSMWISNTATTAMVPIVEALQQMEATSAAATEAGLELVDKGAKELPQSQV 186
Qy 187 VGNKNKNDEPPYPKTYLTCYLYGIAYASSLGGCGCTIIGTATNLTKGIYBARFKNSTE 246
Db 187 IFEGPTLGGQEQBERKRLCKANTLCICYAASIGGTATLTGTGPNVLLGQNNELPDSKD 246
Qy 247 QMDPFTFMEYSVPSMLVYTLTTFVFLQWHFM-----GLWRPKSKEAOEVQREG 296
Db 247 LVNFASFAPAFENMLVMLLPAWLMLQFYMRFNFKSWGCGLESKONEKA----- 297
Qy 297 ADVAKKVIDORYKDLGPMISIHEIQWMLFIWVMVYTRKPGIFLGWADLL-----NSKDI 352
Db 298 ---ALKVLOEYRKLGPLSFABINVLICFFLLVILWFSRDPGMPGWLTVANVEGETKYV 354
Qy 353 RNSMPTIFVVMCFMLPA---NYAFELRYCTRRGGVPVPTGPTPSLITWKPIQTKVPWGLVF 409
Db 355 SDATVAIFVATLLFIVPSQKPKFN-RSQTEERKTPFP-PPLLDWKVTQEKVPWGIVL 412
Qy 410 LGGGFALAEBSQSGMAKLIGNALIGKVLPSVLLLVV-ILVAVELTAFSSNVAIANI 468
Db 413 LGGGFALAKGSEASGLSVWGMKOMEPLHAVPPAAITLILSLVAVF-TECTSNVATTTL 471
Qy 469 IIPVLAEMSLATEIHPLYLILPAGLACSMAPHLPVSTPPNALVAGYANIRTKDMAIAGIG 528
Db 472 FLPIFASMSRSIGLNPYIIMLPCTLSASFAPMLPVPATPPNAIVFTYGHKLVADMVKTGVI 531
Qy 529 PTIIITIILFVFCQWGLVVPVPLNSFPPEWAOI 561
Db 532 MNIIIGVFCVFLAVNTWGRAIF-DLDHFPDQWNV 563
```

Search completed: June 30, 2005, 09:32:46
Job time : 77.2245 secs

THIS PAGE IS BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2005, 08:29:45 ; Search time 17.0834 Seconds
(without alignments)
3221.603 Million cell updates/sec

Title: US-10-017-479A-2

Perfect score: 2994

Sequence: 1 MEIEIGEPQPPVKCSNFFA.....NSFPEWAQIYAAALGNKTH 572

Scoring table: BLOSOM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : PIR 79:*

1: piri:*

2: piri2:*

3: piri3:*

4: piri4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1024	34.2	593	2	146528 sodium/dicarboxyla
2	1003	33.5	568	2	JC7911 Na+-coupled citrat
3	885	29.6	595	2	A47714 Na+/sulfate cotran
4	859	28.7	599	2	T21613 hypothetical prote
5	815	27.2	545	2	T18694 hypothetical prote
6	773	25.8	526	2	S43561 YCR37C homolog K08
7	773	25.8	534	2	G88575 protein K08E5.2 (i
8	759.5	25.4	539	2	A88546 protein R107.1 (im
9	730.5	24.4	539	2	S30871 hypothetical prote
10	676.5	22.6	520	2	C89980 hypothetical prote
11	632	21.1	552	2	F64546 sodium-dependent t
12	626	20.9	548	2	E71961 probable transpor
13	511	17.1	446	2	F70302 transporter (Pho87
14	447	14.9	461	2	I64080 probable membr
15	445.5	14.9	471	2	A81948 probable transmembr
16	444.5	14.8	471	2	G81156 transporter, Nadc
17	440.5	14.7	443	2	C69205 sodium/dicarboxyla
18	426.5	14.2	462	2	B82510 transporter, Nadc
19	411	13.7	456	2	AF0093 Sodium, sulfate sym
20	391.5	13.1	432	2	H64383 Na+ transporter -
21	383	12.8	487	2	C82215 probable transport
22	374	12.5	867	2	T40336 probable MSF trans
23	367	12.3	478	2	A82336 probable transport
24	336.5	11.2	517	2	C89840 conserved hypothet
25	302.5	10.1	894	2	S45135 probable membrane
26	300.5	10.0	513	2	I39534 hypothetical prote
27	281.5	9.4	450	2	A10909 probable membrane
28	277	9.3	881	2	S46633 probable membrane
29	244.5	8.2	923	1	MMBY7C probable membrane

30	212	7.1	425	2	C75019 transport protein
31	209.5	7.0	424	2	F71205 hypothetical prote
32	190	6.3	157	2	S36784 mucin - rat (fragm
33	188.5	6.3	589	2	H84072 sodium/sulfate sym
34	185	6.2	482	2	F71969 hypothetical prote
35	184	6.1	479	1	B64043 conserved hypothet
36	184	6.1	487	1	Q8ECRS YGJ8 protein - Esc
37	179	6.0	618	2	F84409 arsenite transport
38	177.5	5.9	589	2	F75398 transporter, sodiu
39	177	5.9	487	2	B91122 hypothetical prote
40	177	5.9	487	2	A85967 hypothetical prote
41	176.5	5.9	612	2	S74936 sulfur deprivation
42	164.5	5.5	592	2	D85393 hypothetical prote
43	164	5.5	470	2	G86516 dicarboxylase tran
44	162	5.4	470	2	H72105 dicarboxylase tran
45	161	5.4	429	1	B41902 arsenical pump mem
46	160.5	5.4	838	2	S28911 Gene DN10 protein
47	159	5.3	449	2	S76839 hypothetical prote
48	159	5.3	501	2	A50578 citrate carrier (i
49	158	5.3	475	2	A10010 probable membrane
50	155	5.2	471	2	F71543 probable dicarboxy
51	152.5	5.1	411	2	G90154 arsenite transport
52	152.5	5.1	532	2	A57173 ocucutaneous alb
53	150.5	5.0	469	2	T35526 probable integral
54	150	5.0	428	2	G83803 hypothetical prote
55	150	5.0	429	1	C41903 arsenical pump mem
56	148	4.9	430	2	G89962 hypothetical prote
57	148	4.9	610	2	AG0312 probable ion trans
58	146	4.9	477	1	B64813 ybH protein - Esc
59	146	4.9	488	2	B81831 probable integral
60	145	4.8	608	2	A81194 C4-dicarboxylate t
61	145	4.8	430	2	AC0798 probable sodium/su
62	143	4.8	425	2	F72315 conserved hypothet
63	142	4.7	477	2	G85579 probable membrane
64	142	4.7	477	2	F90728 probable membrane
65	142	4.7	487	1	B64795 ybS protein - Esc
66	142	4.7	487	2	G85560 probable membrane
67	142	4.7	487	2	C90710 probable membrane
68	140.5	4.7	436	2	D98100 conserved hypothet
69	140	4.7	585	2	S69216 sulfur deprivation
70	136.5	4.6	478	1	F69811 2-oxoglutarate/mal
71	136.5	4.6	484	2	A50987 probable membrane
72	135	4.5	610	2	H83165 probable sodium/su
73	134.5	4.5	436	2	C95236 membrane protein (
74	134	4.5	590	2	AG3106 transporter, sodiu
75	134	4.5	590	2	E98180 AcrB/AcrD/AcrF fam
76	133.5	4.5	1046	2	AD2959 probable rnd efflu
77	133.5	4.5	1046	2	B98324 probable transport
78	133	4.4	610	2	A85870 probable transport
79	133	4.4	610	2	H91025 probable transport
80	133	4.4	610	2	B65001 probable transport
81	132	4.4	421	2	G72300 conserved hypothet
82	131.5	4.4	428	2	H70528 probable arsB prot
83	131	4.4	620	2	H82761 sulfur deprivation
84	129.5	4.3	759	2	A46188 CAMP-activated Na+
85	128.5	4.3	461	2	AB2185 hypothetical prote
86	127	4.2	489	2	AF2970 conserved hypothet
87	127	4.2	489	2	D98312 hypothetical prote
88	124.5	4.2	447	2	C86130 gluconate transpor
89	124.5	4.2	447	2	H91288 gluconate transpor
90	124.5	4.2	447	2	S56546 gluconate permease
91	120	4.0	431	2	A11599 arsenic efflux pum
92	119	4.0	440	2	G84069 hypothetical prote
93	118	3.9	428	2	E81324 probable arsenical
94	118	3.9	548	2	AC3587 sulfur deprivation
95	116	3.9	430	2	C81402 probable integral
96	116	3.9	495	1	S26870 NADH2 dehydrogenas
97	115	3.8	522	2	A83804 involved in spore
98	115	3.8	527	2	T49900 2-oxoglutarate/mal
99	114.5	3.8	427	2	B83001 probable C4-dicarb
100	114	3.8	372	2	G90740 hypothetical prote

Thu Jun 30 13:13:30 2005

ALIGNMENTS

RESULT 1

I46528
sodium/dicarboxylate cotransporter - rabbit (fragment)
C:Species: Oryctolagus cuniculus (domestic rabbit)
C>Date: 14-Feb-1997 #sequence_revision 14-Feb-1997 #text_change 09-Jul-2004
C:Accession: I46528
R:Paigor, A.M.
J. Biol. Chem. 270, 5779-5785, 1995
A:Title: Sequence and functional characterization of a renal sodium/dicarboxylate cotransporter
A:Reference number: I46528; MUID:95197598; PMID:7890707
A:Accession: I46528
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-593 <PAJ>
A:Cross-references: UNIPROT:Q28615; EMBL:U12186; NID:G758383; PIDN:AAA99666.1; PID:G758383
C:Superfamily: sodium/sulfate cotransporter

Query Match 34.2%; Score 1024; DB 2; Length 593;
Best Local Similarity 37.6%; Pred. No. 5.1e-77;
Matches 219; Conservative 127; Mismatches 186; Indels 50; Gaps 10;

QY 20 ANHWKG-----LVVFLVLLCLPVMMLNEGAEFRCYMLLVMAIFWVTEALPLYVTSM 72
DB 2 ATCQGLWAYRMYLLVFLPLSLPLVPRKEAYCAYALLMALFWCTDALPLAVTAL 61
QY 73 IPIVAFPIGIMSSDQTCRLYFKDTLVNFMGGINVALAVEYCNLHKRLALRVLIQVGCSP 132
DB 62 LPLCLFPWKGMEASEVGLVYKDTNVLFIGLLLAIAVEHNLHKRIALRVLLITGVPR 121
QY 133 RRLHFLGIMVTFMSWISNAACAMMCPPIQVLELO-----AQGVCKINHE----- 181
DB 122 ALLILGFVVTAFSLMSWISNTASTAMVPIAHAVLOELNQTQSNVEGSDNPTELQEPS 181
QY 182 PQYQIVGNGKKNDEPPYPT-----KTLTCVYLGIAVASSLGCGT 223
DB 182 PQKTSKVDEKNDGQAPLPAVPLESGEHTWQELRFSQNSLC-----VCYSASIGGIAT 237
QY 224 IIGTATNLITFKGIYEAREPKNSTEQMDFTFMFVSPMLVYTLTFFVLQWHFGLWRPK 283
DB 238 LTGTPNLVLQGMSTLPPQNPVNFASWFGFAPFIMVILLLSWLMLQLIFLGINFRK 297
QY 284 SKBAQVORGEGADVAKVVDQYKDLGPMSEIHEIQMILFIMVVMYFTRKPGIELGW 343
DB 298 NFGIREQEHQORQAAYRVITQYRLGLGPMSEFAKAVFILLVLLWFTREPGFHW 357
QY 344 ADLNSKD-----IRNSMPTIFVVMCFMPLPANYAFRYCTRGGPVPTGPTPSLITWKF 398
DB 358 GNLVFSASGRVMSDGSASILIGVLFVWPSKIPGLTQDPNFGRLKA--PPALLNWK 415
QY 399 IQTKVPWGLVFLGGGFALAGSKQSGMAKLIQNALIGLK-VLPNSVLLVILVAVFLT 457
DB 416 VNKMPWNLVLLGGYALAKGSESGLSQMLGNKIMPLQHVPPPVATVFIICLLVATF-T 474
QY 458 AFSSNVATANIILPVAEMLSLAIEIHPYLLIIPAGLACSMAFHLPVSTPPNALVAGYANI 517
DB 475 ECTSNAAATLLPLPILASQAQICLHPYLMPLCTLASLAFMLPVATPPNAIVFSFGL 534
QY 518 RYKDMAIAGIGPTIITILFVFCQTWGLVVPNLNSFPFWA 559
DB 535 RYSDMARAGIMLNIIGLVIMLAINSGVPMF-QLHTFFPSWA 575

RESULT 2

JC7911
Na+/sulfate cotransporter NaCT - human
C:Species: Homo sapiens (man)
C>Date: 31-Mar-2003 #sequence_revision 31-Mar-2003 #text_change 14-Jul-2003
C:Accession: JC7911
R:Inoue, K.; Zhuang, L.; Ganapathy, V.
Biochem. Biophys. Res. Commun. 299, 465-471, 2002

A:Title: Human Na+-coupled citrate transporter: Primary structure, genomic organization,
A:Reference number: JC7911; MUID:22334959; PMID:12445824
A:Accession: JC7911
A:Molecule type: mRNA
A:Residues: 1-568 <INO>
A:Cross-references: GB:AV151833
C:Comment: This transporter classified as a tricarboxylate transporter represents the fi:
the cellular entry of citrate by a process energized by the electrochemical Na+ gradient
or the synthesis of fatty acid and chol 17p12-13
C:Genetics:
A:Gene: nact

Query Match 33.5%; Score 1003; DB 2; Length 568;
Best Local Similarity 37.1%; Pred. No. 2.7e-75;
Matches 214; Conservative 124; Mismatches 185; Indels 54; Gaps 13;

QY 19 FANHKGVLV-FLVPLCLCLPVMMLNEGAEFRCYMLLVMAIFWVTEALPLYVTSMPIVA 77
DB 7 YVSKFKSFVLEFVTPLELLPLVILMPAKFVRCAVYIILMAIYWCTEVPLAVTSLMPVL 66
QY 78 FPIGIMSSDQTCRLYFKDTLVNFMGGINVALAVEYCNLHKRLALRVLIQVGCSPRRLHF 137
DB 67 FFLQLDLDRQVCVQYKDTNMLFLGLLIVAVAVERNLHKRIALRVLLVWGAKPRLML 126
QY 138 GLIMVTFMSWISNAACAMMCPPIQVLELOAGVC-----KINHEPQYQI 186
DB 127 GFMGVTFALLSMWISNTATTAMVPIEALQOMEATSAATEAGLELVDRKGAKELPGSQV 186
QY 187 VGG-----NKKNEDEPPYPTKITLCYLGIAVASSIGCGTIIGTATNLITFKGIYEAREPK 242
DB 187 IFEGPILGOQDQERKRLCKAMTLC-----ICVAASIGGTATLTGTGPNVLLIQMNELEP 242
QY 243 NSTEQMDFTFMFVSPMLVYTLTFFVLQWHF-----GLMRPKSKEAQEVOR 292
DB 243 DSKDLVNFASFAPFAPNMLVLLFAWLQFVYMRFNFKSWGCGLESKNEKA----- 297
QY 293 GREGADVAKVIDQYKDLGPMSEIHEIQMILFIMVVMYFTRKPGIFLGWADLL-----N 348
DB 298 -----ALKVQEEYRKLGLPLSFAEINVLICFFLLVILWFSRDPGFMPLTVANVEGE 350
QY 349 SKDIRNSMPTIFVVMCFMPLPA---NYAFRYCTRGGPVPTGPTPSLITWKFIOTKVPW 405
DB 351 TKTVSDATVAIFVATLGFIVPSQKPFNP-RSQTEERKTPFPY-PPLLDMKVYQEKVPW 408
QY 406 GLVFLGGGFALAGSKQSGMAKLIQNALIGLKVLPSNVLVLLV-ILVAVFLTAFSSNVA 464
DB 409 GIVLLGGGFALAGSKSEASGLSVWKGQMEPLHAVPPAAITLLSLVAVF-TECTSNVA 467
QY 465 IANIIPVLAEMSLAIEIHPYLLIIPAGLACSMAFHLPVSTPPNALVAGYANIRTKDMAI 524
DB 468 TTTFLPLFIFASMSRSIGLNPLYIMLPCTLSASFAFMPLPVATPPNALVFFYGHUKVADWVK 527
QY 525 AGIGPTIITILFVFCQTWGLVVPNLNSFPFWAQI 561
DB 528 TGVIMNIIIGVFCVFLAVNTWGRAIF-DLDHFPDWNV 563

RESULT 3

A47714
Na+/sulfate cotransporter, renal - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 02-Jun-1994 #sequence_revision 02-Jun-1994 #text_change 09-Jul-2004
C:Accession: A47714
R:Markovich, D.; Forgo, J.; Stange, G.; Biber, J.; Murer, H.
Proc. Natl. Acad. Sci. U.S.A. 90, 8073-8077, 1993
A:Title: Expression cloning of rat renal Na+/SO4(2-) cotransport.
A:Residues: 1-595 <MAR>
A:Reference number: A47714; MUID:93376745; PMID:7690140
A:Accession: A47714
A:Status: preliminary
A:Molecule type: mRNA

A:Cross-references: UNIPROT:Q07782; GB:L19102; NID:G310182; PIDN:AAA41677.1; PID:G310183
C:Superfamily: sodium/sulfate cotransporter

Query Match 29.6%; Score 885; DB 2; Length 595;
 Best Local Similarity 34.2%; Pred. No. 1.7e-65;
 Matches 202; Conservative 123; Mismatches 190; Indels 76; Gaps 14;

QY 26 LVVFLVPLCLPMLNNEGAEFRCMYLLVLAIFWVTEALPLVYVTSWIPVAPPIGMS 85
 DB 15 LVVFTV-LVLLPLPLIIRSKAEACVILFVIATFTEALPISITALLPLGMPFPGMS 73
 QY 86 SDQTCRLYFKDTLVMFMGIMVALAVEYCNLHKRLALRVIOVGCSPRLHFLGLIWMF 145
 DB 74 STHVASAYFKDHLILLIGVCLATSEKWNLHKRIALRMVMGVNPAWLITGFSSTAP 133
 QY 146 LSWISINAACTAMMCPPIQAVLEEL-----QAQVCKINHEPOYQIV 187
 DB 134 LSWLSNTSTAAMVPIVEAVAQOITSABAEAEATQMTYFNESAQGL-----EVDETII 188
 QY 188 G-GNKKNEDEP-----PYPTK-----ITLCYLG 212
 DB 189 GOETNERKEKTPALGSSNDKGVSSKMETEKNTVTGAKYRSKDHMMCKMLCLC---I 244
 QY 213 AYASSLGGCGTIGTATNLTFFGIYEAERPKNSTEOMDPTFMFYSVPSMLVTLTFFVL 272
 DB 245 AYSSITGGITGTSTNLIFSEHFNTRYD-CRCLNFGSWFLFPFPVAVILLLSWML 303
 QY 273 QWHFMGLMRPKSKAEQVQREGADVAKKVIDQYKIDLPMSIHEIQVMILFIFVMVY 332
 DB 304 QMLFLGFNFKMFCKGKTTLKEKA--CAEVIKQEVKLGPMRYQEIIVLIVFIVMALLW 361
 QY 333 FTRKRGIFLGWADLLNSKD--IRNSMPTIFVVMCFMPLPANYAFRYCTRRGGPVPTGT 390
 DB 362 FSRDFGFTGWSLVFSEYGYVTDSTVALVAGILFLLIPAKK--LTQMTSTGDIIFDYS 419
 QY 391 PSLITWKTQTKVPCGLVPLGGLGFPALARGSKOSGMKILGNALIGLKLVPNSVLLVVI 450
 DB 420 P-LITWKEFQSPMDIALVGGGFPALADGCGVGLSSWIGSKLPLGSLPWLILISS 478
 QY 451 LVAVELTAFSSVAJANIIIPVLAEMSLAIEIHPYLLIPAGLACSMAPFLPVSTPPNAL 510
 DB 479 LIVTSLEVASNPATITLFPILSLPLAEAIHVNPLHILLPSTLCTSFAPLLPVANPPNAI 538
 QY 511 VAGYANIRTKMAIAGIGTITITLTFVFCOTWGLVVP--NLNSFPEWA 559
 DB 539 VFSYGLHKLVIDMVKAGLVNIGLVAVVLMGMFTW---IEPMENLHEYPSWA 586

RESULT 4
 T21613
 hypothetical protein F31F6.6 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
 C:Accession: T21613
 R:Percy, C.
 submitted to the EMBL Data Library, March 1996
 A:Reference number: Z19449
 A:Accession: T21613
 A:Status: preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: DNA
 A:Residues: 1-599 <WIL>
 A:Cross-references: UNIPROT:Q93455; EMBL:Z69884; PIDN:CAA93752.1; GSPDB:GN000028; CESP:F3
 A:Experimental source: clone F31F6
 C:Genetics:
 A:Gene: CESP:F31F6.6
 A:Map position: X
 A:Introns: 39/3; 142/2; 207/2; 255/2; 285/2; 397/1; 437/3; 493/2; 537/3
 C:Superfamily: sodium/sulfate cotransporter

Query Match 28.7%; Score 859; DB 2; Length 599;
 Best Local Similarity 32.6%; Pred. No. 2.5e-63;
 Matches 188; Conservative 128; Mismatches 197; Indels 64; Gaps 11;

QY 23 WKGLVFLVPLCLPMLNNEGAEFRCMYLLVLAIFWVTEALPLVYVTSWIPVAPPI 80

DB 21 WQALLIP-SPLL-----MFVGDHSHGLOAKCLYCVAVMGSVYVFEALPLAITAIPMILFPL 75
 QY 81 MGINSDDTCRLYFKDTLVMFMGIMVALAVEYCNLHKRLALRVIOVGCSPRLHFLGLI 140
 DB 76 FGIMSESEVARAYLPDTCFLFMGGLMVALAVEKCELHARVALFVLKTVGSEPARVMAGPM 135
 QY 141 MYTMPLSMWISINAACTAMMCPPIQAVLELOAQ-----GVCKINHEPOYQIVGNGKK- 192
 DB 136 GVTGFLSMWISINTATTALMVPILQSVITELVSNHRMEDLVALCEAHNHSRKHSGVMRRL 195
 QY 193 -----NNEDE-----PPYPTKITLCYLGIAVASSLGGCGTIGTATNLTFFGIYEA 239
 DB 196 SLPNENNEIKREEMDTAMSPREQKMAKGLMILSVCFSANIGGAATTGTASNLVLVQLNE 255
 QY 240 RPKNSTEOMDPTFMFYSVPSMLVTLTFFLQWHFMGLMRPKSKAEQVQREGADV 299
 DB 256 LPPGADTGVNLSWILFAPF--MVFCCLLYCMCVLYLYL-----RDAPKGSII 302
 QY 300 AKKVIDQYKIDLPMSIHEIQVMILFIFVMVYFTRKPGIFLGWADL----- 346
 DB 303 VTRKLQCKYNELHAFSPAEMAVIFCFALLLVILWILREPVVPGMGEMPKDELVPKSLTEK 362
 QY 347 -----LNSKDIRNSMPTIFVVMCFMPLPANYAFRYCTRRGGPVPTGPTSLITWFIQTK 402
 DB 363 KNTHLTPRFVSDATSAEPIVILLFTLPE-----KLPSRSGSSEQRKASSGLLDWATVQDR 417
 QY 403 VPWGLVFLGGLGFPALARGSKOSGMKILGNALIGLKLVPNSVLLVLAIFWVTEALPFSN 462
 DB 418 FWSVFLGGLGFPALARGSKESGLSHDIGAMRYLDVFNHNIIMLICIIISVTLTNVCSN 477
 QY 463 VAIANIIIPVLAEMSLAIEIHPYLLIPAGLACSMAPFLPVSTPPNALVAGYANIRTKDM 522
 DB 478 TVIASIRIPIVAELARSLEIDPLNFMPLVTTISASFAFLPVATPPNAIVFSSGYLKLVFDM 537
 QY 523 AIAGIGTITITLTFVFCOTWGLVVPNLNSFPEWA 559
 DB 538 FVSGLCVTLGCVLSMLNMLLWAGVFP-NLHLFPQWA 573

RESULT 5
 T18694
 hypothetical protein B0285.6 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
 C:Accession: T18694
 R:Sulston, J.
 submitted to the EMBL Data Library, June 1994
 A:Reference number: Z19007
 A:Accession: T18694
 A:Status: preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: DNA
 A:Residues: 1-545 <WIL>
 A:Cross-references: EMBL:Z34533; PIDN:CAA84299.1; GSPDB:GN000021; CESP:B0285.6
 A:Experimental source: clone B0285
 C:Genetics:
 A:Gene: CESP:B0285.6
 A:Map position: 3
 A:Introns: 6/3; 56/3; 95/2; 107/2; 160/2; 175/1; 211/2; 287/2; 353/1; 393/3; 494/3
 C:Superfamily: sodium/sulfate cotransporter

Query Match 27.2%; Score 815; DB 2; Length 545;
 Best Local Similarity 32.1%; Pred. No. 1e-59;
 Matches 177; Conservative 119; Mismatches 221; Indels 34; Gaps 8;

QY 43 EGAEFRCMYLLVLAIFWVTEALPLVYVTSWIPVAPPIGMSDQTCRLYFKDTLWFM 102
 DB 4 ESTEGKCAFVILTMSCYVAEVPVAVTSFIPMIALPFLGIVSIKEVAPKYFADTNIVFF 63
 QY 103 GGIMVALAVEYCNLHKRLALRVIOVGCSPRLHFLGLIWMVTLMSWISINAACTAMMCPPI 162
 DB 64 NSLMVLAVBECQLHKRIALRMVTVGRPHLWAGMFIITSLTMSIDTACCALMAPI 123
 QY 163 IQAVLEELQAGVCKINHEPOYQIVGNGKNNEDEPPYPTK-----ITLCY 208

Db 124 AYALLEEIM---IPKMRPEKEHEIEVMKIFDKEDPEEKKKLDTSRLSVDRGICKCM 180
Qy 209 YLGIAVASSLGGGCTIIGTATNLTFFKIYEAFKFNSTEQMDFFTFMFYSVPSMLVTLT 268
Db 181 MLLVAHASLIGGTGTINSTGNLIFRDNIEKNFPNEDHGISYLSWMAFAPPMIFWFSS 240
Qy 269 FVFLOW-----HPMGLWR_PKSKEAQEVQREGADVAKKVDQRYKDLGPMSEHIOV 321
Db 241 WFIQLOFLGPRHLGMGRFBETEKOE-----EEVAKRAVMKSYDQGLGPMTWAEKST 293
Qy 322 MLFIEMVMVYFTRKPGIFLGWADLLNSKDIRSMPTIFVVMCFMLPANYAFRLRYCTRR 381
Db 294 LVIFVLAVLSVSDPKVIQFNSDLFRKGVYVUSCSGLVAVFLFWPKKPPDF-IFRK 352
Qy 382 GGPVPTGPTSLITWKFIOQKVPMLVFLGGGFALAEQSKQSGMAKLIGNALIG-LKVL 440
Db 353 DKSRPSVRQBLIDWCVRFRFPWSIILLGAGFAISDAVRVSGLSLIACSLNSTISKM 412
Qy 441 PNSVLLLVILVAVFLTAPSNVAIANIIPVLAEMSLAIEIHPYLILPAGLACSMAPH 500
Db 413 PFFVMOILSVVVVMTFSTNSATSFIPISFKMAEAVGAHPLYSIFTAIGPSFSFM 472
Qy 501 LPVSTPPNALVAGYANIRTKDMAGTGPITITITLTFVFCQTWGLVYVYVNLNSFPWEAQ 560
Db 473 LPMATPANAIYETKTIEMIDMWSGVFLNATFCIAITAINMTWAFWLF-NMGTYPDYAL 531
Qy 561 IYAAALGNKT 571
Db 532 RHATNMTGNS 542

RESULT 6
YCR37C homolog K08E5.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 07-Sep-1994 #sequence_revision 10-Nov-1995 #text_change 24-Nov-1999
C:Accession: S43561
R:Kershaw, J.
submitted to the EMBL Data Library, March 1994
A:Reference number: S43561
A:Accession: S43561
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-526 <KER>
A:Cross-references: EMBL:Z30974
A:Introns: 33/3; 122/2; 169/2; 214/2; 253/2; 481/3
C:Superfamily: sodium/sulfate cotransporter

Query Match 25.8%; Score 773; DB 2; Length 526;
Best Local Similarity 33.2%; Pred. No. 2.9e-56;
Matches 176; Conservative 106; Mismatches 190; Indels 58; Gaps 9;
Qy 46 EPRCYLLVMAIFWVTEALPLYVTSMIPVAPPIMSSDOTCRLYFKDTLVNFMGGI 105
Db 34 EWKCAVCVCIIVYMWSEVPLAVTAMLPVFLPLGVLDANTTAKEYMNDTNFLFGL 93
Qy 106 MVALAVEYCNLHKRLALRVQIVGCSPRRLHGLIMVTFMFLSMWISNAACACTAMCPIQA 165
Db 94 IMAAAVEKCDLHERVALSVLRVCGSEPKWIMLGFTVTALLSSFSINTATTAMVPIQS 153
Qy 166 VLEELQAGVCKINHEPQVQIVGGNKKNEDEPPYPTKITLCYILGIAYASSLGGCGTII 225
Db 154 VVQQL-----ISSFQHP-----TNGERGLCKKQWATGLVLSICFAANIGTGTAT 200
Qy 226 GTATNLTFKGIYEAFKFNSTEQMDFFTFMFYSVPSMLVTLTTFVFLQHPMGLWRPKSK 285
Db 201 GTPSNLVMQLGQSLALFPKVDGSLNYVTWIFAYPLMLLCLFVAMTLLVSFFL-----R 253
Qy 286 EAEQVQREGADVAKKVIDORYKDLGPMSEHIOVMILFIFVMVYFTRKPGIFLGWAD 345
Db 254 DAPEKD-----EAVTEMLKTRYNELPRMTYAEKSVFCFCLLSLWVFRNPVGVPGV 307

Qy 346 LLNSKDIRSMPTIFVVMCFMLP-----ANYAFRLRYCTRRGGPVPTGPTPSLITWKFI 399
Db 308 FFKGAYTDATSAMIVAFLLFVLPSERPDLATYIKKEDLKRG-----CLMDWKTM 358
Qy 400 QTKVPWGLVFLGGGFALAEQSKQSGMAKLIGNALIGLKVLPNSVLLLVILVAVFLTAF 459
Db 359 QETFPMSVLLGGGFALAAAGVKESGLSLLIGNSLSSIEHLPMLWLQLLTLMIAMVITNI 418
Qy 460 SNNVAIANIIPVLAEMSLAIEIHPYLILPAGLACSMAPHLPVSTPPNALVAGYANIR 519
Db 419 CSNTVTASIFVIVATLAQRAGHPHTLMLPTLASSFAFIPFVGTTPNALVFGSGMKV 478
Qy 520 KDMATAGIPTITITITLTFVFCQTWGLVYVYV-----LNSFPWEAQI 561
Db 479 SDMAFVG-----GIISLELLVLT---VLYMNSIAYLTPLLEFPTWAI 519
RESULT 7
G88575
protein K08E5.2 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-May-2001
C:Accession: G88575
R:anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_elegans/
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: G88575
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-534 <STO>
A:Cross-references: GB:chr_III; PIDN:CAA83225.1; PID:g3878357; GSPDB:GN00021; CESP:K08E5.
A:Gene: K08E5.2
A:Map position: 3
C:Superfamily: sodium/sulfate cotransporter

Query Match 25.8%; Score 773; DB 2; Length 534;
Best Local Similarity 33.2%; Pred. No. 3e-56;
Matches 176; Conservative 106; Mismatches 190; Indels 58; Gaps 9;
Qy 46 EPRCYLLVMAIFWVTEALPLYVTSMIPVAPPIMSSDOTCRLYFKDTLVNFMGGI 105
Db 42 EWKCAVCVCIIVYMWSEVPLAVTAMLPVFLPLGVLDANTTAKEYMNDTNFLFGL 101
Qy 106 MVALAVEYCNLHKRLALRVQIVGCSPRRLHGLIMVTFMFLSMWISNAACACTAMCPIQA 165
Db 102 IMAAAVEKCDLHERVALSVLRVCGSEPKWIMLGFTVTALLSSFSINTATTAMVPIQS 161
Qy 166 VLEELQAGVCKINHEPQVQIVGGNKKNEDEPPYPTKITLCYILGIAYASSLGGCGTII 225
Db 162 VVQQL-----ISSFQHP-----TNGERGLCKKQWATGLVLSICFAANIGTGTAT 208
Qy 226 GTATNLTFKGIYEAFKFNSTEQMDFFTFMFYSVPSMLVTLTTFVFLQHPMGLWRPKSK 285
Db 209 GTPSNLVMQLGQSLALFPKVDGSLNYVTWIFAYPLMLLCLFVAMTLLVSFFL-----R 261
Qy 286 EAEQVQREGADVAKKVIDORYKDLGPMSEHIOVMILFIFVMVYFTRKPGIFLGWAD 345
Db 262 DAPEKD-----EAVTEMLKTRYNELPRMTYAEKSVFCFCLLSLWVFRNPVGVPGV 315
Qy 346 LLNSKDIRSMPTIFVVMCFMLP-----ANYAFRLRYCTRRGGPVPTGPTPSLITWKFI 399
Db 316 FFKGAYTDATSAMIVAFLLFVLPSERPDLATYIKKEDLKRG-----CLMDWKTM 366
Qy 400 QTKVPWGLVFLGGGFALAEQSKQSGMAKLIGNALIGLKVLPNSVLLLVILVAVFLTAF 459
Db 367 QETFPMSVLLGGGFALAAAGVKESGLSLLIGNSLSSIEHLPMLWLQLLTLMIAMVITNI 426
Qy 460 SNNVAIANIIPVLAEMSLAIEIHPYLILPAGLACSMAPHLPVSTPPNALVAGYANIR 519

Db 427 CSNTVTASIFVPIVATLQQRAGHPPHTLMLPTTLASSFAFIFFPGTTPNAIVFGSGMWKV 486
 Qy 520 KMAIAGIGFTIITITITIVFVCQTWGLVVPZN-----LNSPPEWAOI 561
 Db 487 SDMAFVG-----GIISLELLVLT---VLVYNSYATLTLPLLEPPTWAIL 527

RESULT 8
 A88546
 protein R107.1 [imported] - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
 C:Accession: A88546
 R:anonymous, The C. elegans Sequencing Consortium.
 Science 282, 2012-2018, 1998
 A:Title: Genome sequence of the nematode C. elegans: a platform for investigation
 A:Reference number: A75000; PMID:93069613; PMID:9851916
 A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projec
 A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103,
 A:Accession: A88546
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-539 <STO>
 A:Cross-references: UNIPROT:P32739; GB:chr_III; PIDN:CAA78468.1; PID:g3879033;
 A:Note: predicted using Genefinder
 C:Genetics:
 A:Gene: R107.1
 A:Map position: 3
 C:Superfamily: sodium/sulfate cotransporter

Query Match 25.4%; Score 759.5; DB 2; Length 539;
 Best Local Similarity 31.8%; Pred. No. 4e-55;
 Matches 177; Conservative 111; Mismatches 205; Indels 63; Gaps 14

Qy 24 KGLVFLVPLCLLPWMLNEGAEPRCMYLLLVMAIFVWTEALPLYVTSMIPVAPPMGI 83
 Db 11 KCLLVLLGLVAVPLLF--GPEYRCLFSIIFLSTYMWIGEAFPIGTSLFPLALYPILOI 68
 Qy 84 MSSDOTCLRYFKDVLVPMFGIMVALAVEYCNLHKRLALRVIQIVGCSPPRLHFGILMTV 143
 Db 69 VFSKQISPVYFKDSIVLFCMLTIVMAVVEATGLHRRRLKLLTKVGAQP----- 118
 Qy 144 MFLSMWISNAACATMCPPIIOAVLEEL-OAQGVCKINHE----POYQIVGGNKKNNEDEP 198
 Db 119 --VSFFVSDTACTALMCPTAVALLMSMSDAVQHLKEDHRKPKPPDDATVAEKLRIIDMT 176
 Qy 199 PYPYTKITLCYVIGIYASSLGGCTIIGTATNLTFKGIYEAPFNKSTOEOMDPTFMFYSV 258
 Db 177 PQDAGFCFALILACAHASLIGTAAITSTGNLVFRENHKKRYPEGQVTMTYLOMWFVFI 236
 Qy 259 PSMLVYVTLTLTFVFLQWPMFG-----LW--RPKSEAEQVQREGADVAKKVIDQR--- 307
 Db 237 PMFVYLLASYYILVCYFMGSTFARFEPSPKEAH-----LKKIEKNIQTM 285
 Qy 308 KYDLGPMSTHETQVMILFIWMVYFTRKPGIFLGMDLLNSKD--IRNSMPTIFVVMVCF 366
 Db 286 YEDLDGVSWGKESVPVPFILLIGSWISRDPGFTPGMGDLLPHRFISDSVSGVLISCLIF 345
 Qy 367 MLPANYAFLYCYTRRGGPV-PTGTPPSLITWKFIQTQKVPWGLVFLGGGFALAEBSKQSG 425
 Db 346 VMPKO-----PFDPIDPMAPILKWTDMKSKFSWCSCTLLICAGYAISEGVKSG 393
 Qy 426 MAKLLIG---NALIGLKVLPSVLLVVLVAVPLTAFSSNVAIANIIPVLAEMSLAIE 481
 Db 394 LSRLLSCGMKNI FVGMSSLP---LQLFTVTIIVIMTEFASNVSTGSIPIPSIGVAESWG 450
 Qy 482 IHPLYLILPAGLACSMAPHLVPSTPPNALVAGYANIRTKDMAIGTPIITIIITLTFVFC 541
 Db 451 VHPLYALPTTVACSPAPMLPISTPPNAVYDTKVISWEMIVCGFELLNACILITSLNM 510
 Qy 542 QTWGLVYVYNLNSRPE 557

Df

S11 NTWTYFIF-SLNIPPE 525

RESULT 9
S30871
hypothetical protein R107.1 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 21-Jan-2000
C/Accession: S30871
R;Thomas, K.
submitted to the EMBL Data Library, July 1992
A:Reference number: S30871
A:Accession: S30871
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-539 <THO>
A:CROSS-references: EMBL:Z14092
C;Genetics:
A:Genetic code: SGCA
A:Introns: 30/3; 80/3; 119/2; 217/2; 251/3; 293/2; 390/3; 416/1; 445/3; 491/3
C/Superfamily: sodium/sulfate cotransporter
C;Keywords: transmembrane protein

Query Match 24.4% Score 730.5; DB 2; Length 539;
Best Local Similarity 30.9%; Pred. No. 1e-52;
Matches 170; Conservative 114; Mismatches 214; Indels 53; Gaps 12

Qy 24 KGLVFLVLPLCLPVMLNEGAPRCMYLLLVMAIFWTEALPLYVTSMIVAPPIMGI 83
| | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 11 KKLVLLGPIVAVPLEFF--GPEYSCLFSMMFLTSTYWIGEAFPIGVTSLFPLALYPILQI 68
| | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

Qy 84 MSSDOTRLYPKDTLVNFMGGIMVALAVEYNLHKRLARVIQVGCSPRHLHFGLINVT 143
| | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 69 VPSKQISPVFKDSWLPMCTLIWMAVEATGLHRSIALKULLTKVGAQP----- 118
| | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

Qy 144 MFLSMITSNAACTAMMCPIIOAVLEEL-QAQVCCKINHE----PQYQIVGNKKQNDEP 198
| | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 119 --VSFSDTACTALMCPATAVALLMSDSA VOHLKEDSHSKPKPPDDATVAEKLSDMDMT 176
| | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

Qy 199 PYPTKITLYVLGITAYASSLGCGGTII GTATNLTFKG IYEARKFNSTBQMDFPTMFYSV 258
| | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 177 PQDAGFCKAULACAHASLHGTTAITSTGP NLVFRENIHKSPYEGQVTMTYLQMWVFM 236
| | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

Qy 259 PSLMVLTLLTVFLQWHFMG-----LWRPKSEAEQVRREGADVAKKVIDQRKYDLG 312
| | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 237 PMPFVYLLASYIILVCYFGMPSTFA RFESPSKEAHLKK-----LIEKNIQTMWE DLG 290
| | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

Qy 313 PMSIHEIQWMLFI FVMVWYTRKPGI GLPWADLL-NSKDIRNSMPI FVVVMCFMLPAN 371
| | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 291 DVSNGEKSFVEVFILLIGSWISHDPGFT PGWGDLPHSHNFMSDSVSGVLISCILFWVPKD 350
| | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

Qy 372 YAFLRVCTRGGGV-P TGPPTSILTWKFIOTKP WGLVFLGGLGFALAEAGSQSGMAXLI 430
| | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 351 -----FPDPIDPA PILKKTWDKSKFSWSCTLL IGAGYAISEGVDKSLSLI 398
| | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

Qy 431 G-----NALIGLKVL PNSVLLVLVAVFUT AFSNNVAIANII IPVLAEMSLAIETHPLY 486
| | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 399 SCGMKRI FVGMSSLP--- LQTVTTI VMTEFASNVS STGISFIP ISLGVAESMGVHPLY 455
| | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

Qy 487 LI LPAGLACSMAPHLPVST PPNALVAGYANIR TKDMAI AGIGPTITI ITTLFVFCQTWGL 546
| | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 456 LA LPTTVACSFAFMLDISTEPNAVVDTKVISWEM MVCCFLNLIA CMLITS LNKNWTY 515
| | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

Qy 547 VVYPNLSNPFE 557
| | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 516 FIF-SLNIPPE 525
| | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

RESULT 10
C89980
hypothetical protein SAL732 [imported] - Staphylococcus aureus (strain N315)
C;Species: Staphylococcus aureus
C;Date: 10-May-2001 #sequence revision 10-May-2001 #text change 09-Jul-2004

C;Accession: C89980
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogino, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kato, C.; Sekimizu, K.; C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A;Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.
A;Reference number: A89758; MUID:21311952; PMID:11418146
A;Accession: C89980
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-520 <KUR>
A;Cross-references: UNIPROT:Q99SX1; GB:BA000018; PID:g13701709; PIDN:BA043002.1; GSPDB:G
A;Experimental source: strain N315
C;Genetics:
A;Gene: SA1732
C;Superfamily: sodium/sulfate cotransporter

Query Match 22.6%; Score 676.5; DB 2; Length 520;
Best Local Similarity 29.9%; Pred. No. 2.9e-48;
Matches 164; Conservative 118; Mismatches 176; Indels 91; Gaps 14;

QY 18 FFANHWKLVVFLVPLCLPMLL--NEGAFRCMYLLV---MAIFWTEALPLVVTSM 72
DB 28 FSACQLICLI--LGPLLFLTLFFHPQDLFWKGVVYLAITLWATWITEAIPATSL 85

QY 73 IPIVAFPIGIMSSDOTCRLYFKDTLVFMGGMVALAVEYCNLHKLRLRVIOVGCSP 132
DB 86 LPVILLPLGHILTPQVSSEYNDIIFELGGFILAEMRWLHTRVALTIINLGAST 145

QY 133 RRLHFLGIMVTFMFLMWSNAACTAMCPITIOAVLELOAQGVCKINHPEQYIOVGNKK 192
DB 146 SKILLGFVATGFLSMFVSNTAAVMIMPIGLAIKEA-----HDLQ-----E 188

QY 193 NNEDEPPYPTKITCYVILGTAAYASLGGCTIIGTATNLTFKGIYEAFKNSTEQMDFFT 252
DB 189 ANTQST--QKFEKSLVLAIGYAGTIGGLTGLTLPILILKGOYMQHFGH---EISFAK 244

QY 253 FMFYSVPSMLVYLLTFVFLQW-----HFMGLWRPKSKEAQVQREGADVAKKVI 304
DB 245 WMIVGIPITVILLGITWILYKVAFRHDLKYLPG-----GQTLI 283

QY 305 DQRYKDLGPMSEIHIQVMILFIEMVVMYFTRKPGIFLGWADLNSKDIRNSMP-----TIF 360
DB 284 KQKLDLGMKYEKVVQTIPLVASLLMITRE-----FLKKWEVTSVADGTTAIF 335

QY 361 VVMCFMPLPANYAFELVCTRGGVPPTPSLITWKFIQTKVPWGLVFLGGFALAEG 420
DB 336 ISILLFIIPAK-----NTEKHRIIDWE--VAKELPWGVLLIFGGGLALAKG 380

QY 421 SKQSGMAKLIGNALIGLKVLPNSVLLAVILVAVFLTAFSSNVAIANIIPVLAEMSLAI 480
DB 381 ISEGLAKWLGEQLKSLNGVSPILIVITVFLTEVTSNTATATMILPILATLSVAV 440

QY 481 EIHPYLLTLPAGLACSMFAHLPVSTPPNALVAGVANTRTKDMATAGIPITIIITLFPV 540
DB 441 GVHPLLMAPAAMAANCAYMLPVGTTPNALIFGSGKISIKQMASVGFVWNLISAIIII-- 498

QY 541 CQTWGLVYV 549
DB 499 -----LVVY 502

RESULT 11
F64546
sodium-dependent transporter - Helicobacter pylori (strain 26695)
C;Species: Helicobacter pylori
C;Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 09-Jul-2004
C;Accession: F64546
R;Tomb, J.F.; White, O.; Kervatage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenne-son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.; Nature 388, 539-547, 1997
A;Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.

A;Title: The complete genome sequence of the gastric pathogen *Helicobacter pylori*.
A;Reference number: A64520; MUID:97394467; PMID:9252185
A;Accession: F64546
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-552 <TOM>
A;Cross-references: UNIPROT:Q25003; GB:AE000541; GB:AE000511; NID:g2313299; PIDN:AA00728;
C;Superfamily: sodium/sulfate cotransporter

Query Match 21.1%; Score 632; DB 2; Length 552;

Best Local Similarity 31.9%; Pred. No. 1.5e-44;
Matches 164; Conservative 94; Mismatches 186; Indels 70; Gaps 14;

QY 54 LVMAIFWVTEALPLVVTSMIPVAPPIGIMSSDOTCRLYFKDTLVFMGGMVALAVEY 113

DB 72 VLMGIWMTTEALDLPATALLPLVFSVSDQFASVSSYASPIIFLFGGFIILSMQK 131

QY 114 CNLHKLRLRVIOVGCSPRRLHFLGIMVTFMFLMWSNAACTAMCPITIOAVLELOAQ 173

DB 132 WNLHTRIALSIILLVGTSPRRLILGFMATGFLSMVSNATATAVMLPVGMSVL-QLVAK 190

QY 174 GYCKIN-----HEPO--YQIVGNKKN-----NEDEPPYPTKITCYVILGIA 214

DB 191 LVKEDASNSWQKEEIKAHGIMSNIVHKGDITQVIOEKTIVRTNFSICLMGLIA 250

QY 215 ASSLGGCGTIICTATNLTFKGIYEAFKNSTEQMDPPTFMFYSVPSMLVYLLTFVFLQW 274

DB 251 AASIGSLGTLTTPNALLAGYMKTAFA---NIEIDFAQWVFGTPLAFTMLILAMLLTY 307

QY 275 HFVGLWRPKSKEAQVQREGADVAKKVIDORYKDLGPMSEIHIQVMILFIEMVVMY 334

DB 308 VIFPL-----KIKTPGKE-----VIRVELKKGLRSQAEISVGIIILASLGM-- 352

QY 335 RKPGIFLG-----WADLLNSKDIRNSMPTIFVVMCFMPLPANYAFELRYCTRGGVP 389

DB 353 -----IFLGMWLSWGVKIDKID---SVIANGVSALLFILPANHQGR----- 392

QY 390 TPLSLITWFKIQTVPWGLVFLGGGFALAEQSGMAKLIGNALIGLKVLPNSVLLVY 449

DB 393 ---LIDWG--VAKKLPWDVLLFGGLALSAQFSKTLGSLWIGLHNSGFSHLFILFIVMV 448

QY 450 ILVAVELTAFSSNVAIANIIPVLAEMSLAIEH-PLVLLIPAGLACSMFAHLPVSTP 506

DB 449 TLMVIELTEITNTATAAFLPVIGGVANGMVGHENQSLLTIPVALSATCAFMFPVTP 508

QY 507 PNALVAGVANTRTKDMATAGIGPTIITITLFPV 540

DB 509 PNAIAGSGVVKITDMIKAGLMNLVGVVLISTF 542

RESULT 12

E71961

probable transporter - Helicobacter pylori (strain J99)

C;Species: Helicobacter pylori

A;Variety: strain J99

C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004

C;Accession: E71961

R;Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;

Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;

Nature 397, 176-180, 1999

A;Title: Genomic sequence comparison of two unrelated isolates of the human gastric path

A;Reference number: A71800; MUID:99120557; PMID:9923682

A;Accession: E71961

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-548 <ARN>

A;Cross-references: UNIPROT:Q9ZML8; GB:AE001458; GB:AE001439; NID:g4154713; PIDN:AA005784;

A;Experimental source: strain J99

C;Genetics:

A;Gene: jhp0200

C;Superfamily: sodium/sulfate cotransporter

Query Match

20.9%; Score 626; DB 2; Length 548;

Qy	145	FLSMISNAACTAMCPPIQAVLBELOAQGVCKINHBEPQYQIVGGNKKNNEDEPPPTTKI	204
Db	126	LISMWISNTSTTTAMMLPALGIL-----AGV-----RETEREKVPFPVL	164
Qy	205	TLCYVLGTAYASSLGGCGTITGATNLTFKGIVYEARFKNSTEQMDPRTFMYFVPSMLVY	264
Db	165	-----LGTAYSASVGIGITLVGSPNGIAGIL-----GLSPFDMLKFGIPVFLIL	210
Qy	265	TLLTFVFLQHFPMGLWRPKSK-EAQEVQREGADAVAKVIDQRYKDLGPMSIHEIQVMI	323
Db	211	FPULLFAIL-----FL-VERPSTDLKVERVQ-----EIKFEPTQR-----VLVLL	249
Qy	324	LIFPFWVMY-FTRKPGIPLGWADLLNSKDIIPNSMPTIFVVMCFMLPANYAFILRYCTR	382
Db	250	IFLFTALAWIFSK-----IAPFEVKVKYFDFTVALLAVVALEFIF-----	289
Qy	383	GPVPTGPTPSLITWKFIOQYPMGLVFLGGGFALABSGSKOSGMAKLIGNALIG-LKVLP	441
Db	290	-----RLLDWRDVKEGVSWGTLILLFGGIALSGIMKKTGTAKFISQELVDVLHGLP	340
Qy	442	NSVLLLVILVAVFIATPSSNAVAIANIIPVLAEMSLAIEIHPLYLILPAGLACSMAPHL	501
Db	341	TELELLITVLFWIELTELMSNTATTALLIAPILFSTAQMIKGPPEMLVIVIPAAVASCA	400
Qy	502	PVSTPPNALVAGYANIRTKDMAIAGIGPTIITILFVF	540
Db	401	VPATPENAIVGTGYIKQSQMMRVGLINIVFSIVLAAF	439

RESULT 14

I64080

probable membrane protein H10608 - Haemophilus influenzae (strain Rd KW20)

C/Species: Haemophilus influenzae

C/Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 09-Jul-2004

C/Accession: I64080

R/;Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Ker
D.M.; Brandon, R.C.; Fink, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.;
Gocayne, J.D.; Scott, J.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen,
Science 269, 498-512, 1995

A/Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Frazer, C.M.; Smith, H.O.

A/Title: Whole-genome random sequencing and assembly of Haemophilus influenzae

A/Reference number: A64000; MUID:95350630; PMID:7542800

A/Accession: I64080

A/Status: nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-461 <TIGR>

C/Cross-references: UNIPROT:Q57486; GB:U32743; GB:I42023; NID:g1573597; PIDN:A

C/Superfamily: probable transporter M0672

Query Match	14.9%	Score 447;	DB 2;	Length 461;
Best Local Similarity	24.2%	Pred. No. 2.8e-29;		
Matches 127;	Conservative 115;	Mismatches 178;	Indels 104;	Gaps 16
QY	26	LVVFLVPLLCUPV-MLNLGEAFRCNYLLLVMAIFWVTEALPLYVTS-MIPIVAFPMIGI	83	
Db	20	IVLFPVLLNVLPFPKANSG-----LALLAFIAVLMLSEALHTVITALLVPLIAV-ALGL	73	
QY	84	HSSDQTCRLYFKDTLVFMFGIMVALAVECNLHKRLALRVIOIVCSPRRHLFGILMVT	143	
Db	74	VSTQALVGFADPTIFLFFCGFSLATALTQKDKLIANKMALARGNLFIAYIVYLFIT	133	
QY	144	WFLSMWISNAACTAMWCPIIQAVLEELQAGVCKINHBPQYIVGGNKKNDEPPYPTK	203	
Db	134	APLSMWNENTATAMVPLPLANGILSUDRE-----KQHTNY-----	169	
QY	204	ITLCYVLGIAYASSLGGCGTIIGTATNLTPKGIYEAFKNSTEQMDPFTFMFVSPSMLV	263	
Db	170	-VVFLLGIAYSAGISGWTGLVGSPPNAIVA-----SNLNTFSDLWGLVGPIMII	218	
QY	264	YTLTFTFVLQHFHGLWRPKSKBAQEVQREGADVAKKVIDORYKDLGFGMSIHEIQWMI	323	
Db	219	LLPLMIGLYIT-----IFPKK-----LHLNFQETPNI---ENNPMRLIT	255	

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2005, 08:28:39 ; Search time 76.9736 Seconds
(without alignments)
3805.323 Million cell updates/sec

Title: US-10-017-479A-2

Perfect score: 2994

Sequence: 1 MEIEIGEQPPVKCSNPFA.....NSFPWAQIYAAALGNKTH 572

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 100 summaries

Database : UniProt 03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2394	100.0	572	1	IND1_DROME	Q9vvt2 drosophila
2	2394	100.0	590	2	Q7KUS6	Q7kus6 drosophila
3	1618.5	54.1	896	2	Q7QIT2	Q7qit2 anopheles g
4	1454.5	48.6	504	2	Q86B89	Q86b89 drosophila
5	1427	47.7	505	2	Q811B6	Q811b6 drosophila
6	1382.5	46.2	562	1	IND2_DROME	Q9vvdq drosophila
7	1271	42.5	414	2	Q7PF67	Q7pf67 anopheles g
8	1042.5	34.8	587	2	Q35055	Q35055 rattus norv
9	1033.5	34.5	586	1	S132_MOUSE	Q9es88 mus musculu
10	1028	34.3	586	2	Q6AZR9	Q6azr9 xenopus lae
11	1026	34.3	592	1	S132_HUMAN	Q11183 homo sapien
12	1024.5	34.2	624	2	Q642N5	Q642n5 xenopus tro
13	1024	34.2	593	1	S132_RABIT	Q28615 oryctolagus
14	1023.5	34.2	613	2	Q803K7	Q803k7 brachydanio
15	1022.5	34.2	613	2	Q6P3L2	Q6p3l2 brachydanio
16	1010.5	33.8	622	2	Q57661	Q57661 xenopus lae
17	1007	33.6	605	2	Q866R0	Q866r0 didelphis m
18	1004	33.5	568	2	Q86YTS	Q86yts homo sapien
19	1002	33.5	568	2	Q6ZMGI	Q6zmg1 homo sapien
20	993.5	33.2	587	1	S132_RAT	P70545 rattus norv
21	983	32.8	572	2	Q8CJ44	Q8cj44 rattus norv
22	975.5	32.6	600	1	S133_RAT	Q9z025 rattus norv
23	972.5	32.5	600	1	S133_MOUSE	Q9ly63 mus musculu
24	965.5	32.2	602	1	S133_HUMAN	Q8wvt9 homo sapien
25	964.5	32.2	602	2	Q8IVB1	Q8ivb1 homo sapien
26	948.5	31.7	572	2	Q67BT3	Q67bt3 mus musculu
27	910	30.4	596	2	Q6NWY4	Q6nwy4 brachydanio
28	901.5	30.1	601	2	Q7ZWL2	Q7zwl2 xenopus lae
29	893	29.8	602	2	Q6DJN7	Q6djn7 xenopus lae
30	892.5	29.8	594	1	S131_MOUSE	Q9jhi4 mus musculu
31	890.5	29.7	594	2	Q80YB5	Q80yb5 mus musculu

32	889.5	29.7	627	1	S134_HUMAN	Q9ukg4 homo sapien
33	888	29.7	602	2	Q6E7G8	Q6e7g8 xenopus lae
34	885	29.6	595	1	S131_RAT	Q07782 rattus norv
35	885	29.6	626	2	Q8N631	Q8n631 homo sapien
36	883.5	29.5	625	2	Q8BZ82	Q8bz82 mus musculu
37	879.5	29.4	595	1	S131_HUMAN	Q9bzw2 homo sapien
38	878.5	29.3	583	2	Q6NY28	Q6ny28 brachydanio
39	878.5	29.3	583	2	Q6PE27	Q6pe27 brachydanio
40	876	29.2	601	2	Q6W1I2	Q9w712 pseudopleur
41	874.5	29.2	582	1	NAD1_CAEEL	Q93655 caenorthabdi
42	838.5	28.0	619	2	Q6L970	Q6l970 anguilla ja
43	819	27.4	520	2	Q8N8K4	Q8n8k4 homo sapien
44	812	27.1	577	1	YK66_CAEEL	P46556 caenorthabdi
45	799.5	26.7	566	1	NAD3_CAEEL	Q21339 caenorthabdi
46	791.5	26.4	551	1	NAD2_CAEEL	P32739 caenorthabdi
47	695	23.2	552	2	Q8NEE3	Q8ene3 oceanobacil
48	687	22.9	546	2	Q65NCO	Q65nc0 bacillus li
49	676.5	22.6	520	2	Q8NV55	Q8nv55 staphylococ
50	676.5	22.6	520	2	Q99SX1	Q99sx1 staphylococ
51	676.5	22.6	520	2	Q7A4P8	Q7a4p8 staphylococ
52	676.5	22.6	520	2	Q6G816	Q6g816 staphylococ
53	673.5	22.5	520	2	Q6GFE0	Q6gfe0 staphylococ
54	644	21.5	626	2	Q6AQ77	Q6aq77 desulfotale
55	632	21.1	552	2	Q25003	Q25003 helicobacte
56	626	20.9	548	2	Q9ZML8	Q9zml8 helicobacte
57	621.5	20.8	540	2	Q8LG88	Q8lgh8 arabidopsis
58	620.5	20.7	450	2	Q75MH3	Q75mh3 homo sapien
59	619.5	20.7	540	2	Q9MAW4	Q9maw4 arabidopsis
60	611	20.4	510	2	Q7UUK9	Q7uuk9 rhodospirell
61	610.5	20.4	510	2	Q8NTS7	Q8nts7 corynebacte
62	610.5	20.4	527	2	Q6M8D0	Q6m8d0 corynebacte
63	609.5	20.4	524	2	Q6NK59	Q6nk59 corynebacte
64	607.5	20.3	495	2	Q67T81	Q67t81 symbiobacte
65	607.5	20.3	540	2	Q9FEH9	Q9feh9 arabidopsis
66	602	20.1	462	2	Q9FGK1	Q9fgk1 arabidopsis
67	601.5	20.1	474	2	Q55MM0	Q55mm0 bacillus li
68	596	19.9	587	2	Q8FU27	Q8fu27 corynebacte
69	586.5	19.6	548	2	Q6ZFH7	Q6zfh7 oryza sativ
70	580	19.4	450	2	Q8FMZ5	Q8fmz5 corynebacte
71	528.5	17.7	302	2	Q68D44	Q68d44 homo sapien
72	511	17.1	446	2	Q66449	Q66449 aquifex aeo
73	485.5	15.5	463	2	Q9CLQ0	Q9clq0 pasteurella
74	459	15.3	459	2	Q8T3Y2	Q8t3y2 drosophila
75	450.5	15.0	470	2	Q65R88	Q65r88 manheimia
76	447	14.9	461	1	Y608_HAEIN	Q57486 haemophilus
77	446	14.9	456	2	Q6LQW4	Q6lqm4 photobacter
78	445.5	14.9	471	2	Q9JV43	Q9jv43 neisseria m
79	444.5	14.8	471	2	Q9K032	Q9k032 neisseria m
80	440.5	14.7	443	2	Q26881	Q26881 methanobact
81	431.5	14.4	461	2	Q7N822	Q7n822 photorhabdu
82	430.5	14.4	456	2	Q87Q94	Q87q94 vibrio para
83	430.5	14.4	474	2	Q8GF70	Q8gf70 photorhabdu
84	426.5	14.2	462	2	Q9KNE0	Q9kne0 vibrio chol
85	424	14.2	459	2	Q66615	Q66615 yersinia ps
86	420.5	14.0	456	2	Q7MLG6	Q7mlg6 vibrio vuln
87	420.5	14.0	456	2	Q8D908	Q8d908 vibrio vuln
88	420	13.9	451	2	Q7M8Y1	Q7m8y1 wolinnella s
89	417	13.9	459	2	Q8CZT3	Q8czt3 yersinia pe
90	411	13.7	456	2	Q8ZHX0	Q8zhx0 yersinia pe
91	411	13.7	463	2	Q65SS5	Q65ss5 manheimia
92	409.5	13.7	464	2	Q9AED9	Q9aed9 leptospira
93	409.5	13.7	464	2	Q72QJ5	Q72qj5 leptospira
94	409	13.7	464	2	Q8VTX1	Q8vtx1 leptospira
95	405.5	13.5	464	2	Q8F5L4	Q8f5l4 leptospira
96	403.5	13.5	464	2	Q9S4F3	Q9s4f3 leptospira
97	400	13.4	464	2	Q9ZGJ3	Q9zgj3 leptospira
98	391.5	13.1	432	1	Y672_METJA	Q58086 methanococc
99	390.5	13.0	471	2	Q6LSU9	Q6lsu9 photobacter
100	383	12.8	487	2	Q9KSE3	Q9ksa3 vibrio chol

ALIGNMENTS

RESULT 1
ID IND1 DROME STANDARD; PRT; 572 AA.
AC Q9VV72; Q9NHY9;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 25-JAN-2005 (Rel. 46, Last annotation update)
DE I'm not dead yet protein (INDY transporter protein) (drIndy).
GN Names=Indy; ORFNames=CG3979;
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophiliidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RN SEQUENCE FROM N.A., AND FUNCTION.
RP MEDLINE=22250702; PubMed=12186628; DOI=10.1042/BJ20021132;
RX Inoue K., Fei Y.-J., Huang W., Zhuang L., Chen Z., Ganapathy V.;
RA "Functional identity of Drosophila melanogaster Indy as a cation-
RT independent, electroneutral transporter for tricarboxylic acid-cycle
RT intermediates.";
RL Biochem. J. 367:313-319 (2002).
RN [2]
RN SEQUENCE FROM N.A.
RP STRAIN=Berkely;
RX MEDLINE=2019606; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Anantides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers J.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Franckoch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Deicher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foslter C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laeko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleeb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamis I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195 (2000).
RN [3]
RN GENOME REANNOTATION.
RP MEDLINE=22426069; PubMed=12537572;
RX Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,

RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review.";
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22 (2002).
RN [4]
RN SEQUENCE FROM N.A.
RP STRAIN=Berkely; TISSUE=Embryo;
RX MEDLINE=22426066; PubMed=12537569;
RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,
RA George R.A., Guarin H., Krommiller B., Pacleeb J.M., Park S., Wan K.H.,
RA Rubin G.M., Celniker S.E.;
RT "A Drosophila full-length cDNA resource.";
RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8 (2002).
RN [5]
RN SEQUENCE OF 366-572 FROM N.A.
RP STRAIN=Canton-S;
RX MEDLINE=20050669; PubMed=10581279;
RA Schmid K.J., Nigro L., Aquadro C.F., Tautz D.;
RT "Large number of replacement polymorphisms in rapidly evolving genes
RT of Drosophila. Implications for genome-wide surveys of DNA
RT polymorphism.";
RL Genetics 153:1717-1729 (1999).
RN [6]
RN IDENTIFICATION, FUNCTION, AND TISSUE SPECIFICITY.
RP PubMed=1118146; DOI=10.1126/science.290.5499.2137;
RA Rogina B., Reenan R.A., Nilsen S.P., Helfand S.L.;
RT "Extended life-span conferred by cotransporter gene mutations in
RT Drosophila.";
RL Science 290:2137-2140 (2000).
RN [7]
RN FUNCTION, SUBCELLULAR LOCATION, INDUCTION, AND TISSUE SPECIFICITY.
RP PubMed=1231301; DOI=10.1073/pnas.222531899;
RX Knauf F., Rogina B., Jiang Z., Aronson P.S., Helfand S.L.;
RA "Functional characterization and immunolocalization of the transporter
RT encoded by the life-extending gene Indy.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14315-14319 (2002).
RN -1- FUNCTION: Cation-independent electroneutral transporter (not
CC associated with membrane depolarization) of a variety of
CC tricarboxylic and dicarboxylic acid-cycle intermediates. There is
CC also small, but detectable, transport of monocarboxylics.
CC Transport is through the epithelium of the gut and across the
CC plasma membranes of organs involved in intermediary metabolism and
CC storage. Affinity for substrates is citrate > succinate >
CC pyruvate. Fumarate, a-ketoglutarate, and glutamate are also
CC transported, but not lactate. Transport mechanism that is not
CC coupled to Na(+), K(+), or Cl(-). Function is shown in Xenopus
CC oocytes and human retinal pigment epithelial (HRPE) cell lines.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein; basolateral
CC membrane of cells in the midgut.
CC -1- TISSUE SPECIFICITY: In adults, abundantly expressed in the fat
CC body, basolateral region of midgut cells and oenocytes. Low level
CC expression is seen in the halteres, procardia, restricted regions
CC of the esophagus and hindgut, base of the legs and in a subset of
CC cells in the third segment of the antennae.
CC -1- INDUCTION: Completely inhibited by DIDS. Modest but significant
CC inhibition by phloretin or furosemide.
CC -1- MISCELLANEOUS: The life-extending effect of mutations is likely
CC caused by an alteration in energy balance caused by a decrease in
CC transport function.
CC -1- SIMILARITY: Belongs to the SLC13A transporter family. NACD
CC subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; AF509505; AAN86815.1; -;
CC EMBL; AE003519; AAF49226.1; -;
DR

```
DR EMBL: AV102686; AAM27515.1; --
DR EMBL: AF217399; AAF73384.1; AUT_INIT.
DR IntAct: Q9VVT2; --
DR GO: FlyBase; FBgn0036816; Indy.
DR GO: GO:0005887; C:integral to plasma membrane; IDA.
DR GO: GO:0015137; F:citrate transporter activity; IDA.
DR GO: GO:0005833; F:pyruvate transporter activity; IDA.
DR GO: GO:0015141; F:succinate transporter activity; IDA.
DR GO: GO:0015746; P:citrate transport; IDA.
DR GO: GO:0008340; P:determination of adult life span; IMP.
DR GO: GO:0006848; P:pyruvate transport; IDA.
DR GO: GO:0015744; P:succinate transport; IDA.
DR InterPro: IPR001898; Na/sul_symport.
DR Pfam: PF00939; Na_sulph_sym; 1.
DR PROSITE: PS01271; NA_SULFATE; 1.
KW Transmembrane; Transport.
FT TRANSMEM 25 45 Potential.
FT TRANSMEM 64 84 Potential.
FT TRANSMEM 97 117 Potential.
FT TRANSMEM 135 155 Potential.
FT TRANSMEM 206 226 Potential.
FT TRANSMEM 252 272 Potential.
FT TRANSMEM 311 331 Potential.
FT TRANSMEM 355 375 Potential.
FT TRANSMEM 439 459 Potential.
FT TRANSMEM 491 511 Potential.
FT TRANSMEM 522 542 Potential.
SQ SEQUENCE 572 AA; 63552 MW; C021D77D006C5444 CRC64;

Query Match 100.0%; Score 2994; DB 1; Length 572;
Best Local Similarity 100.0%; Pred. No. 2.7e-216;
Matches 572; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEIEIGEOPQPVKCSNFANHWKGLVFLVPLLCPLVWMLNEGAEFRCMYLLVWYAFW 60
DB 1 MEIEIGEOPQPVKCSNFANHWKGLVFLVPLLCPLVWMLNEGAEFRCMYLLVWYAFW 60
QY 61 VTEALPLVYVTSIMPIVAPFIMGIMSSDQTCRLYFKDTLVMFMGGIMVALAVEYCNLHKRL 120
DB 61 VTEALPLVYVTSIMPIVAPFIMGIMSSDQTCRLYFKDTLVMFMGGIMVALAVEYCNLHKRL 120
QY 121 ALRVQIVGCSRRRLHFLGIMVTFMFLSMWISNAACTAMWCPIIQAVLBELOAQGVCKINH 180
DB 121 ALRVQIVGCSRRRLHFLGIMVTFMFLSMWISNAACTAMWCPIIQAVLBELOAQGVCKINH 180
QY 181 EPQYQIVGSKNNKNEDEPPYPTKITLCYVLGAYASSLGGCGTITGTATNLTFTKGIYEAR 240
DB 181 EPQYQIVGSKNNKNEDEPPYPTKITLCYVLGAYASSLGGCGTITGTATNLTFTKGIYEAR 240
QY 241 FKNSTEQMDPFTFMFYSVPSMLVYVTLTFLVFLQHFHFMGLWRPKSKEAQEVQREGADVA 300
DB 241 FKNSTEQMDPFTFMFYSVPSMLVYVTLTFLVFLQHFHFMGLWRPKSKEAQEVQREGADVA 300
QY 301 KKVIDQRYKDLGPMISHEIQVILFIMVYFTFKPGIFLGWADLLNSKDIRNSMPTIF 360
DB 301 KKVIDQRYKDLGPMISHEIQVILFIMVYFTFKPGIFLGWADLLNSKDIRNSMPTIF 360
QY 361 VVVMCFMLPANYAFRLRYCTRRGGPVPTGPTSLITWKEIOTKVPWGLVFLGGGFALAE 420
DB 361 VVVMCFMLPANYAFRLRYCTRRGGPVPTGPTSLITWKEIOTKVPWGLVFLGGGFALAE 420
QY 421 SKQSGMAKLIGNALIGLKVLPNSVLLVVLVAVFLTAFSSNVAIANIIPVLAEMSLAI 480
DB 421 SKQSGMAKLIGNALIGLKVLPNSVLLVVLVAVFLTAFSSNVAIANIIPVLAEMSLAI 480
QY 481 EIHPLYLILPAGLACSMFHLPVSTPPNALVAGYANIRTKMAIAGIPTITITLFPVF 540
DB 481 EIHPLYLILPAGLACSMFHLPVSTPPNALVAGYANIRTKMAIAGIPTITITLFPVF 540
QY 541 COTWGLVVPYVNLNSPPEWAIYAAALGNKTH 572
DB 541 COTWGLVVPYVNLNSPPEWAIYAAALGNKTH 572

RESULT 2
Q7KUS6 PRELIMINARY; PRT; 590 AA.
ID Q7KUS6;
AC Q7KUS6;
DT 05-JUL-2004 (TRENBLrel. 27, Created)
DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
DE CG3979-PB.
GN Name=Indy; ORFNames=CG3979;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
SEQUENCE FROM N.A.
RP MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RX Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA April J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Bencos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Butris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doul L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foalier C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattai B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kimms I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodgett, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195 (2000).
RN [2]
SEQUENCE FROM N.A.
RP MEDLINE=22426065; PubMed=12537568;
RX Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
RA Pacle B.J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila
RL melanogaster euchromatic genome sequence."
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079 (2002).
RN [3]
SEQUENCE FROM N.A.
RP MEDLINE=22426070; PubMed=12537573;
RX Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
```

RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
a genomic perspective.";
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084 (2002).
RN [4]
RN SEQUENCE FROM N.A.
RX MEDLINE=22426069; PubMed=12537572;
RA Milera S., Croesby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
systematic review.";
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083 (2002).
RN [5]
RN SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RN SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
RN EMBL; AE003519; AAF49227.2; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006814; P:sodium ion transport; IEA.
DR InterPro; IPR001898; Na/sul. symport.
DR Pfam; PF00939; Na sulph. sympt; 1.
DR PROSITE; PS01271; NA_SULFATE; UNKNOWN 1.
SQ SEQUENCE 590 AA; 65585 MW; ECA7857BB6B34414 CRC64;

Query Match 100.0%; Score 2994; DB 2; Length 590;
Best Local Similarity 100.0%; Pred. No. 2.8e-216;
Matches 572; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEIEIGROPPVKCSNFFAHNKGIVLVFLVLLCLPVMLLNGAEFRCMYLLVMAIFW 60
DB 19 MEIEIGROPPVKCSNFFAHNKGIVLVFLVLLCLPVMLLNGAEFRCMYLLVMAIFW 78

QY 61 VTEALPLYVTSMPIVAFPIGIMSSDQTCRLYFKDPLVMFMGIMVALAVEYCNLHKL 120
DB 79 VTEALPLYVTSMPIVAFPIGIMSSDQTCRLYFKDPLVMFMGIMVALAVEYCNLHKL 138

QY 121 ALRVIQVCGSPRRHLFGLIMVTFJSMWISNAACATAMCPITQAVLEELQAGVCKINH 180
DB 139 ALRVIQVCGSPRRHLFGLIMVTFJSMWISNAACATAMCPITQAVLEELQAGVCKINH 198

QY 181 EPQYQIVGGNKKNEDEPPYPTKITLCYLGIAVASSLGCGGTIGTATNLTFGIYEAR 240
DB 199 EPQYQIVGGNKKNEDEPPYPTKITLCYLGIAVASSLGCGGTIGTATNLTFGIYEAR 258

QY 241 FKNSTEQMDPTTFMFYSVPMSLVYLLTFVFLQWHFMGLRPKSKAEQVQREGADVA 300
DB 259 FKNSTEQMDPTTFMFYSVPMSLVYLLTFVFLQWHFMGLRPKSKAEQVQREGADVA 318

QY 301 KKVIDQRYKDLGPMNSIHEIQMILFIPMVVMYFTRKPGIFLGWADLLNSKDINSMTPI 360
DB 319 KKVIDQRYKDLGPMNSIHEIQMILFIPMVVMYFTRKPGIFLGWADLLNSKDINSMTPI 378

QY 361 VVWCMFLPANYAFRLYCTRGRGVPVPTPSLITWKFIQTKVPWGLVFLGGGFPALAE 420
DB 379 VVWCMFLPANYAFRLYCTRGRGVPVPTPSLITWKFIQTKVPWGLVFLGGGFPALAE 438

QY 421 SKOSGMKLIQNALIGLKVLPNSVLLLVILVAVFLTAFSSNVAIANIIPVLAEMSLAI 480
DB 439 SKOSGMKLIQNALIGLKVLPNSVLLLVILVAVFLTAFSSNVAIANIIPVLAEMSLAI 498

QY 481 EIHPLYLILPAGLACSMAPHLPVSTPPNALVAGVANIRTKDMAIGPTIITILFVF 540
DB 499 EIHPLYLILPAGLACSMAPHLPVSTPPNALVAGVANIRTKDMAIGPTIITILFVF 558

QY 541 CQTWGLVVYPNLNSRPDEWAIYVAAAALGNKTH 572
DB 559 CQTWGLVVYPNLNSRPDEWAIYVAAAALGNKTH 590

RESULT 3
QYQIT2 PRELIMINARY; PRT; 896 AA.
ID QYQIT2 PRELIMINARY; PRT; 896 AA.
AC QYQIT2;
DT 01-MAR-2004 (TEMBLrel. 26, Created)
DT 01-MAR-2004 (TEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DE EbiP7623 (Fragment).
GN Name=ebiG7623; ORFNames=ENSANGG000000005753;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB01008807; EAA03999.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006814; P:sodium ion transport; IEA.
DR InterPro; IPR001898; Na/sul. symport.
DR Pfam; PF00939; Na sulph. sympt; 4.
DR PROSITE; PS01271; NA_SULFATE; UNKNOWN 1.
FT NON_TER 1
FT NON_TER 896
SQ SEQUENCE 896 AA; 98460 MW; D59FAE475FEE7982 CRC64;

Query Match 54.1%; Score 1618.5; DB 2; Length 896;
Best Local Similarity 55.6%; Pred. No. 7.1e-113;
Matches 298; Conservative 102; Mismatches 111; Indels 25; Gaps 3;

QY 24 KGIWVFLPCLLPVMLNEGAEFRCMYLLVMAIFWVTEALPLYVTSMPIVAFPIGMI 83
DB 386 QSLAVLLVPLIAASVFLVDTSPDFRCMWVWVGWVTEALPLTITSMPLVFLPLMGI 445

QY 84 MSDQTCRLYFKDPLVMFMGIMVALAVEYCNLHKLALRVIOVCGSPRRHLFGLIMVT 143
DB 446 LDTNRTCMYMKDITMLFMFIGGIVMALAVEYCNLHKLVALKVISVWGCSQRRLTFGLTMVT 505

QY 144 MFLSMWISNAACATAMCPITQAVLEELQAGVCKINHPEQYQIVGGNKKNEDEPPYPTK 203
DB 506 MFLSMWISNAACATAMCPITQAVLEELQAGVCKINHPEQYQIVGGNKKNEDEPPYPTK 563

QY 204 ITLCYLYLGIAVASSLGCGGTIGTATNLTFGIYEARFKNSTEQMDPTTFMFYSVPMSLV 263
DB 564 ITWCYIYGTAATAATCGVGTIVSGVNLTKGIYESRFPQA-PGIDFSPFMFNIPIGMILL 622

QY 264 YTLTTFVFLQWHFMGLRPKSKAEQVQREGADVAKKVIDORYKDLGPMNSIHEIQMIL 323
DB 623 FTFTLWMLQWLFMGLFRPNSDARAANTIGPEGEVAERVIENRLELGMTSHEIGVAF 682

QY 324 LFIFMVVMYFTRKPGIFLGWADLLNSKDINSMTPIFVWCMFLPANYAFRLYCTRGRG 383
DB 683 FFVLSVVLFTTREGFMGMWADLVDPVKIKDATPALFVIMLFIIVPADWNCLAFFHKGKG 742

QY 384 PVPTGPTPSLITWKFIQTKVPWGLVFLGGGFPALAEKSGOSGMKLIQNALIGLKVLPNS 443
DB 743 -----NLLFLGGGFPALAEKSGRTSGMSALIGQSLAGLKVLPPL 780

QY 444 VLLLVILVAVFLTAFSSNVAIANIIPVLAEMSLAIETHPLYLILPAGLACSMAPHLPV 503
DB 781 LLLFLVLCVVGEGFTEFTSTNVAICNVVLPVLAETIETHPLVMPVAMWCSFSFHLPV 840

QY 504 STPNALVAGYANIRKDMAGIAGTITITITLTVFCQWGLVUVVNLSPPEWA 559
 Db 841 GTPPNAIVAGVGNIRKDMAGIAGTITITLTVFCQWGLVUVVNLSPPEWA 896

RESULT 4
 Q86B89 PRELIMINARY; PRT; 504 AA.

AC Q86B89;
 DT 01-JUN-2003 (T-EMBLrel. 24, Created)
 DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
 DE CG32921-PB (CG32921-pb)
 GN ORFNames=CG32921;
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]

SEQUENCE FROM N.A.
 RP MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
 RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA Sutton G.G., Wortman J.R., Richards S., Ashburner M., Henderson S.N.,
 RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
 RA Abril J.F., Aghayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadiou E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
 RA Fiesler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glaeser K.,
 RA Glöck A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lascko P., Lei Y., Levitsky A.C., Li J., Li Z., Liang X., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milchena N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner E., Venter E., Wang A.H., Wang X.,
 RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodger, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhou S., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*."
 RL Science 287:2185-2195 (2000).
 RN [2]

SEQUENCE FROM N.A.
 RP MEDLINE=22426065; PubMed=12537568;
 RA Celnik S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,
 RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
 RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
 RA Pacle J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
 RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
 RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
 RT "Finishing a whole-genome shotgun: Release 3 of the *Drosophila*
 melanogaster euchromatic genome sequence."
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079 (2002).

RN [3]
 RP MEDLINE=22426070; PubMed=12537573;
 RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J., Svirskas R.,
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
 RA Ashburner M., Celnik S.E.;
 RT "The transposable elements of the *Drosophila melanogaster* euchromatin:
 a genomics perspective."
 RL Genome Biol. 3:RESEARCH0084-RESEARCH0084 (2002).
 RN [4]

SEQUENCE FROM N.A.
 RP MEDLINE=22426069; PubMed=12537572;
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 RA Battencourt B.R., Celnik S.E., de Grey A.D., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.;
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a
 systematic review."
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083 (2002).
 RN [5]

SEQUENCE FROM N.A.
 RP FlyBase;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 RN [6]

SEQUENCE FROM N.A.
 RP FlyBase;
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE003728; AAC41577.1; -
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005215; P:transporter activity; IEA.
 DR GO; GO:0006814; P:sodium ion transport; IEA.
 DR InterPro; IPR001898; Na/sul_sympot.
 DR Pfam; PF00939; Na_sulph_symp; 2.
 SQ SEQUENCE 504 AA; 5386 MW; 28FDBFCE65065F4A CRC64;

Query Match 48.6%; Score 1454.5; DB 2; Length 504;
 Best Local Similarity 52.4%; Pred. No. 8e-101;
 Matches 268; Conservative 99; Mismatches 135; Indels 9; Gaps 4;

QY 50 MYLLVMAIFWVTEALPLYVTSMIPVAFPTMGIMSSDQTCRLYFKDTLVNFMGIMVAL 109
 Db 1 MYLVANMALFWITEALPLYVTSMIPVAFPTMGIMSSDQTCRLYFKDTLVNFMGIMVAL 60

QY 110 AVECNLHRLALRIQVGCSPRLHFLGIMVTFMSWISNACTAMCPIQAVLE 169
 Db 61 AIEYNLHRLALRIQVGCSPRLHFLGIMVTFMSWISNACTAMCPIQAVLE 120

QY 170 LQAGVCKINHEPOVQIVGKKNEDPPVPTKITLCYVLGIAYVASSLGGCGTIGTAT 229
 Db 121 METQNIFAIYKQEEPV-----EGDPPHPTISMAFYGIAYVSSIGCGTIGTAT 174

QY 230 NLTFKGIYEARFNKSTQMDFPTFMFYSVP-SMLVYVTLTLLTFVQLQWHFMGLWRPKSKEAQ 288
 Db 175 NLTYKGLYDTRFNSDEKIDFPIFMAYSVPVPLVLFITVFSQVTHMGLFRNSKIGQ 234

QY 289 EVQGRGADVAKVVIQORYKIDQPMISHEIQVLMIFVWVYFTRKPGIFLGWADLLN 348
 Db 235 EVKGAESQDVVKDVIQKRAELGPMSCHEIQVLMIFVWVYFTRKPGIFLGWADFLN 294

QY 349 SKDIRNSMPTIFVVMCMFLPANYAFURYCTRRGVPVPTGPTSLITWKEITQKVPWGLV 408
 Db 295 AXAIGSGPPVFFATILLPALPTQVTFPKYCGK-APPFGQTLACLSSVYCKYTPFGLA 353

QY 409 FLLGGGFALAEKSGQSGMAKLIAGLKLKLPNSVLLVVLVAFVLTAFSSNVAIANI 468
 Db 354 FLL-GGFALAEKSGVSGMAKLIAGLKLKLPNSVLLVVLVAFVLTAFSSNVAIANI 412

QY 469 IIPVLAEMSLAIEIHPYLYLILPAGLACSMAPHLPVSTPPNALVAGYANIRKDMAGIAG 528
 Db 413 LIPIFSEMALAIEVHPMKLTFPPALACSLAFHLPVSTPPNALVAGYANIRKDMAGIAG 472

```
QY 529 PTIITITLTFVFCQWGLVVPVNLSPPEWA 559
DB 473 PTCWAFICLLFTGTGTMTLITPCTGTEPPSWA 503

RESULT 5
Q811B6 PRELIMINARY; PRT; 505 AA.
ID Q811B6
AC Q811B6;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Indy-2-PA.
GN Name=Indy-2;
OS Drosophila pseudoobscura (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7237;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tucson 14011-0121.4;
RA Bergman C.M., Pfeiffer B.D., Rincon-Limas D.E., Hoskins R.A.,
RA Gierke A., Mungall C.J., Wang A.M., Kronmiller B., Pacleb J., Park S.,
RA Stapleton M., Wan K., George R.A., de Jong P.J., Botas J., Rubin G.M.,
RA Celniker S.E.;
RT "Assessing the impact of comparative genomic sequence data on the
RT functional annotation of the Drosophila genome."
RL Genome Biol. 3:research0086-research0086(2002).
DR EMBL: AY190949; AAC01056.1; -
DR FlyBase: FBgn0064420; DpseIndy-2.
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0005215; F:transporter activity; IEA.
DR GO: GO:0006814; P:sodium ion transport; IEA.
DR InterPro: IPR001395; Aldo/ket red.
DR Dfam: PF00939; Na sulph symd; 2.
DR PROSITE: PS00063; ALDOXETO_REDUCTASE 3; UNKNOWN 1.
SQ SEQUENCE 505 AA; 55608 MW; C1F8E3D37C579A40_CRC64;

Query Match 47.78; Score 1427; DB 2; Length 505;
Best Local Similarity 52.48; Pred. No. 9.3e-99;
Matches 271; Conservative

QY 50 MYLLVMAIFWTEALPLYTSMIPVAFPIGIMSSDQCRLYFKDKTLVMPGGINVAL 109
DB 1 MYLVSNMALFWTEAIPLYLTSLFPVVFPLFDILGSDQVCKLYFSDTVMVFGGLIAL 60
QY 110 AVEYCNLHRLALRVIOVGCSPRLHFLGLIMVTMFLSMWISNAACATMMCPITIOAVLEE 169
DB 61 AIEYVSNLHQRAMKTLILVGCSPRLHFLGLIMVTMFLSMWISNAATMMCPIVKAVLSE 120
QY 170 LQAGVCVKI----NHPEQYQIVGKNKNNEDEPPYPTKITLCYVLGIAYASSLGGCGTII 225
DB 121 LDSQNIITVYKSEEPW-----EGDPHPSTISMAFYGVAYAAATIGGCGTLI 170
QY 226 GTATNLTFKGIYARFNKNSTQMDPPTMFYSPVSMVLVYLLTFVF-LQWHFGLMRPKS 284
DB 171 GTGTNLTFKGLYDRFRPKSKTQIDFIFEMAYAIPIVVIVNVNLLYFSLQVTHMGLFRGKS 230
QY 285 KEAEVORHGREGADVAKVVDORVKDLPMSIHBIQVWMLFIFVMVYFTRKPGIFLGWA 344
DB 231 QTGLEVRKGTGQAVVTIVIKARHQLQPMTCHEIQVTVILFVLMVFLFRKFGFVTGNG 290
QY 345 DLLNSKDIRNSMPTIFVVMCMFLPANYAFIRYCTREGPVPTGPT-PSLITWKFIQTKV 403
DB 291 DFLNAQKIGSGPPVWLPVMLLFAIPTQYTFKYCC--GKPFPTQSQMDALLSWIYIHRNT 348
QY 404 PWGLVFLGGGFALAEKSGKMAKLGKLVLPNSVILLVLTAVFLTAFFSSNV 463
DB 349 PWGLCFL-LGGFALAEKSGKMAKLGKLVLPNSVILLVLTAVFLTAFFSSNV 407
```

```
QY 464 ATANIIIPVLAEMSLAIEHPLYLILPAGLACSMAPFLPVSTPPNALVAGYANIRTKDMA 523
DB 408 AICNLILPISFEMALAKWHLKLTLPSPSLAISMAVHLPVSTPPNAIISYAGIKTKYLA 467
QY 524 IAGIGTITITITLTFVFCQWGLVVPVNLSPPEWAQ 560
DB 468 LAGILPTIAFWIVLNLNSQ-YGNIIPPESTKFPDWAQ 503

RESULT 6
IND2_DROME
ID IND2_DROME STANDARD; PRT; 562 AA.
AC Q9VDQ0; Q95T83;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 25-JAN-2005 (Rel. 46, Last annotation update)
DE I'm not dead yet protein 2.
GN Name=Indy-2;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkelley;
RA MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.P., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostel C., Gabrielian A.B., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector R., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
RN [2]
RP GENOME REANNOTATION.
RA MEDLINE=22426069; PubMed=12537572;
RX Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
```


Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M., Lewis S.E.;
"Annotation of the Drosophila melanogaster euchromatic genome: a systematic review";
Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
[3]
SEQUENCE FROM N.A.
STRAIN=Berkeley; TISSUE=Head, and Testis;
MEDLINE=22426066; PubMed=12537069;
Stapleton M., Carlson J.W., Brostein P., Yu C., Champe M.,
George R.A., Guarin H., Konniller B., Pacleb J.M., Park S., Wan K.H.,
Rubin G.M., Celnik S.E.;
"A Drosophila full-length cDNA resource.";
Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).
[4]
IDENTIFICATION.
PubMed=1118146; DOI=10.1126/science.290.5499.2137;
Rogina B., Reenan R.A., Nilsen S.P., Helfand S.L.;
"Extended life-span conferred by cotransporter gene mutations in Drosophila.";
Science 290:2137-2140(2000).
CC -I- FUNCTION: Cation-independent electroneutral transporter (not associated with membrane depolarization) of a variety of tricarboxylic and dicarboxylic acid-cycle intermediates. There is also small, but detectable, transport of monocarboxylates. Transport is through the epithelium of the gut and across the plasma membranes of organs involved in intermediary metabolism and storage (By similarity).
CC -I- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -I- SIMILARITY: Belongs to the SLC13A transporter family. NADC subfamily.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AF003728; AAN1391.1; -
EMBL; AF060287; AAL25326.1; -
EMBL; BT001329; AAN71084.1; -
FlyBase; FBgn0011603; Indy-2.
GO; GO:0005887; C:integral to plasma membrane; ISS.
GO; GO:0015137; F:citrate transporter activity; ISS.
GO; GO:0050833; F:pyruvate transporter activity; ISS.
GO; GO:0015141; F:succinate transporter; ISS.
GO; GO:0015746; P:citrate transport; ISS.
GO; GO:0008340; P:determination of adult life span; ISS.
GO; GO:0006848; P:pyruvate transport; ISS.
GO; GO:0015744; P:succinate transport; ISS.
InterPro; IPR001898; Na/sul_symp.
Pfam; PF00939; Na_sulph_symp; 1.
PROSITE; PS01271; NA_SULFATE; FALSE_NEG.
KW Transmembrane; Transport.
FT TRANSMEM 24 44 Potential.
FT FT 64 84 Potential.
FT TRANSMEM 93 113 Potential.
FT TRANSMEM 135 155 Potential.
FT TRANSMEM 200 220 Potential.
FT TRANSMEM 249 269 Potential.
FT TRANSMEM 308 328 Potential.
FT TRANSMEM 350 370 Potential.
FT TRANSMEM 393 413 Potential.
FT TRANSMEM 430 450 Potential.
FT TRANSMEM 478 498 Potential.
FT TRANSMEM 515 535 Potential.
SQ SEQUENCE 562 AA; 61934 MW; D9F04C05FAA73ED3 CRC64;

Query Match 46.2%; Score 1382.5; DB 1; Length 562;
Best Local Similarity 46.5%; Pred. No. 2.3e-95;
Matches 262; Conservative 120; Mismatches 164; Indels 17; Gaps 7;

PROSITE; PS01271; NA_SULFATE; UNKNOWN 1.	
Q SEQUENCE 414 AA; 45265 MW; 800E40E0BBAA168F CRC64;	
Query Match	42.5%; Score 1271; DB 2; Length 414;

Best Local Similarity 57.6%; Pred. No. 3.9e-87;
Matches 235; Conservative 73; Mismatches 92; Indels 8; Gaps 3;
QY 158 MMCPITQAVLEELQAOQCKINHE-----POYQIVGNGKKNQDEPPYTKITLCYVLGI 212
Db 1 MMCPIMQAVLEELQAOQCKINHE-----KEKMDDDTPRPTKTLCLYFLGA 58
QY 213 AYASSLGCCCTIIGTATNLTKGIYEFARFKNQTEQWDFPTFMFYSVPSMLVYTLTLFVEL 272
Db 59 AYASTLGGCCCTIIGTATNLTKGIYEFARFKNQTEQWDFPTFMFYSVPSMLVYTLTLFVEL 117
QY 273 QWFMGLWRKPSKEAQRQVGRGADVAKKVIDQRYKDLGPMSEIHEIQVMILFIFVMVY 332
Db 118 QWLYMGFRNPSPEAKADGAGEAVAKVETRYELGPMSEIHEIQVMILFIFVMVY 177
QY 333 FTRKPGIFLGWADLLNSKDIRNSMPTIFVVMCFMPLPANYAFURYCTRGGVPVPTGTPS 392
Db 178 FTRPGFITGWADLVNTVKIKDATPAMFIVIVLFIPIANWRCLKFCRSNPGRLPTATPG 237
QY 393 LITWKFTQTKVPCVLVLLGGGALAGSKQSGWAKLIGNALIGLKLPLNSVLLVILV 452
Db 238 LITWKFINQKVPWSLIFLLGGGALAGSKQSGWAKLIGNALIGLKLPLNSVLLVILV 297
QY 453 AVFLTAFSSNVAIITIPVLAEMSLAIEIHPYLIILPAGLACSMAPFLPVSTPPNALVA 512
Db 298 AQLTETFSNVAICNVLPVLAEMSLAIEIHPYLIILPAGLACSMAPFLPVSTPPNALVA 357
QY 513 GYANIRTKQMAIAGIGTITITLTFVFCQWGLVVPVNLNSPFWAQ 560
Db 358 GVGNAIKDVAAGIGPSITFLVIMASFPPTGWAVVYPELATPFDWAR 405

RESULT 8

O35055 PRELIMINARY; PRT; 587 AA.
AC O35055;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Sodium/dicarboxylate cotransporter 1 (NA(+)/dicarboxylate
DE cotransporter 1) (Kidney dicarboxylate transporter) (SDCT1) (Organic
DE anion transporter 1) (OAT1).
GN Name=Nadc-1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=KIDNEY;
RX MEDLINE=98355911; PubMed=9691021;
RA Sekine T., Cha S.H., Hosoyamada M., Kanai Y., Watanabe N., Furuta Y.,
RA Fukuda K., Igarashi T., Endou H.;
RT "Cloning, functional characterization, and localization of a rat renal
RT Na(+)-dicarboxylate transporter.";
RL Am. J. Physiol. 275:F298-F305(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=KIDNEY;
RX MEDLINE=97373539; PubMed=9228014; DOI=10.1074/jbc.272.30.18526;
RA Sekine T., Watanabe N., Hosoyamada M., Kanai Y., Endou H.;
RT "Expression cloning and characterization of a novel multispecific
RT organic anion transporter.";
RL J. Biol. Chem. 273:20972-20981(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY;
RX MEDLINE=98361968; PubMed=9694847; DOI=10.1074/jbc.273.33.20972;
RA Chen X.Z., Shayakul C., Berger U.V., Tian W., Hediger M.A.;
RT "Characterization of a rat Na(+)-dicarboxylate cotransporter.";
RL J. Biol. Chem. 273:20972-20981(1998).
CC -!- FUNCTION: MAY MEDIATE ELECTROGENIC, SODIUM-DEPENDENT TRANSPORT OF
CC MOST KREBS CYCLE INTERMEDIATES, INCLUDING CITRATE, SUCCINATE,

CC ALPHA-KETOGLUTARATE, AND OXALOACETATE. TRANSPORT OF CITRATE WAS PH
CC SENSITIVE. IT MAY HAVE A SODIUM:SUBSTRATE STOICHIOMETRY OF 3:1. AT
CC LEAST ONE NA(+) SEEMS TO BIND TO THE TRANSPORTER BEFORE THE
CC SUBSTRATE.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -!- TISSUE SPECIFICITY: EXPRESSED IN S3 SEGMENTS OF KIDNEY PROXIMAL
CC TUBULES AND IN ENTEROCYTES LINING THE INTESTINAL VILLI. ALSO
CC DETECTED IN LUNG BRONCHIOLI, THE EPIDIDYMIS, AND LIVER.
CC EMBL: AB001321; BAA28609.1; -;
DR EMBL: AF058714; AAC31165.1; -;
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0015293; F:symporter activity; IEA.
DR GO: GO:0005215; F:transporter activity; IEA.
DR GO: GO:0006814; P:sodium ion transport; IEA.
DR GO: GO:0006810; P:transport; IEA.
DR InterPro: IPR001898; Na/sul_symport.
DR Pfam: PF00939; Na_sulph_sym; 1.
DR TIGRFAM: TIGR00785; dase; 1.
DR PROSITE: PS01271; NA_SULFATE; 1.
KW Sodium transport; Symport; Transmembrane; Transport.
FT TRANSMEM 13 73 POTENTIAL.
FT TRANSMEM 53 73 POTENTIAL.
FT TRANSMEM 80 100 POTENTIAL.
FT TRANSMEM 114 134 POTENTIAL.
FT DOMAIN 123 126 POLY-LEU.
FT TRANSMEM 215 235 POTENTIAL.
FT TRANSMEM 264 284 POTENTIAL.
FT DOMAIN 272 275 POLY-LEU.
FT TRANSMEM 319 339 POTENTIAL.
FT TRANSMEM 367 387 POTENTIAL.
FT TRANSMEM 445 465 POTENTIAL.
FT TRANSMEM 477 497 POTENTIAL.
FT TRANSMEM 506 526 POTENTIAL.
FT TRANSMEM 535 555 POTENTIAL.
FT CARBOHYD 570 570 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 580 580 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 587 AA; 64142 MW; 84847CT778F9CDA CRC64;
Query Match 34.8%; Score 1042.5; DB 2; Length 587;
Best Local Similarity 38.1%; Pred. No. 7.8e-70;
Matches 219; Conservative 129; Mismatches 190; Indels 37; Gaps 8;
QY 26 LVVFLVPLLCPLVMLNEGAFCMYLLVMAIFWTEALPLVYVTSMIPIVAPINGIMS 85
Db 15 LVLCPIELPLPLIVQTKAYCAYSIIMALLWCTEALPLAVTALFVLPFLMGIMD 74
QY 86 SDQTCRLYFKDVLVPMFGIMVALAVEYCNLHKRLALRVIOVGCSPRRHLHGMVTF 145
Db 75 ASEVCIEYFKDNLVFGGLVMAIAVEHNLHKRIALQVLLIIGVFPALLLGLFVLTA 134
QY 146 LSMWISNAACAMCPIIOAVLEELQA--QGVCKINHEPQYQIV-----GGNK 191
Db 135 LSMWISNTATTAMVPIGHAVLEQLQSKKDVGGNNNPTFELQECQKEVTKLNDGQP 194
QY 192 KNNEDP-----PYPTKITLYYLGTAYASSLGGCGTIIIGTATNLTFKGIYAREK 242
Db 195 VSAPSEPTQKTQEHHRFSQGLSLC-----ICYSASIGGIATLTGTPLNVLQGVNSLFP 250
QY 243 NSTEQMDPPTFMFYSVPSMLVYTLTLFVFLQWFMGLWRKPSKEAQRQVGRGADVAKK 302
Db 251 QNGNVNFAWFGFAPFTMIILLLAWLQVLFVGNFRKNFGFGEGERKQA--AFQ 308
QY 303 VIDQRYKDLGPMSEIHEIQVMILFIFVMVYFTRKPGIFLGWADLLNSKDIRNSMPT---- 358
Db 309 VIKTYRLLGPMSPFAEKTIVTLFVLLVWLWFTREPFGFFPGWGTVFAANEKQSGMSADGV 368
QY 359 -IFVWVNCFMPLPANYAFURYCTRGGVPVPTGPTPSLITWKFIOTKVPWGLVLLGGG 417
Db 369 AIFISLVMFIPSPKIPGLMODPKPKGLKA--PPAILTWKTVDNDKMPNIVILLGGG 426
QY 418 AEGSKQSGMAKLGNALIGLKLPLNSVLLVILVAVFLTAFSSNVAIITIPVLAEMS 477
Db 427 AKGSEQSGLEWLGDKLTPLQHPPPSATAVILCLLIITFAICTESNVAITTLFLPLASMA 486

QY 478 LAIEIHLPLVLLPAGLACSMFHLVSTPPNALVAGYANIRTKDMAIAGIGPTITITL 537
 DB 487 QAICLHPLYVMLPCTLAASLAFMLPVATPPNAIVSFGLKVKVSDMARAGFLNIIIGVLAI 546

QY 538 FVFCOTWGLVVPNLNRPPEMAQIYAAALGNKTH 572
 DB 547 TLSINSWSIPIF-KLDTFPPSWAHSNTSQCLLPSN 580

RESULT 9
 S132 MOUSE
 ID S132 MOUSE STANDARD; PRT; 586 AA.
 AC Q9ES68;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Solute carrier family 13, member 2 (Renal sodium/dicarboxylate cotransporter) (Na(+)/dicarboxylate cotransporter 1) (NADC-1).
 GN Name=S132; Synonyms=Nadcl, Sdct1;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=20425391; PubMed=10966927;
 RA Pajor A.M., Sun N.N.,
 RT "Molecular cloning, chromosomal organization and functional characterization of a sodium/dicarboxylate cotransporter from mouse kidney";
 RL Am. J. Physiol. 279:F482-F490 (2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Colon;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Shenmen C.M., Schuler G.D., Klausner R.D., Collins F.S., Wagner L., Buetow K.H., Schaefer C.F., Bhat N.K., Altchul S.F., Zeeberg B., Moore T., Max S.I., Wang J., Hsieh F., Hopkins R.F., Jordan A., Farmer A.A., Rubin G.M., Hong L., Diatchenko L., Marusina K., Casavant T.L., Schetz T.E., Stapleton M.J., Soares M.B., Bonaldo M.P., Carninci P., Prange C., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Mullahy S.J., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J., Boeak S.A., McEwan P.C., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettner M., Madan A., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.N., Krzywinski M.T., Skalska U., Smalhus D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 CC -!- FUNCTION: Cotransport of sodium ions and dicarboxylates such as succinate and citrate.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -!- TISSUE SPECIFICITY: Highly expressed in kidney and small intestine. Not detectable in brain, heart, stomach and skeletal muscle.
 CC -!- SIMILARITY: Belongs to the SLC13A transporter family. NADC subfamily.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC DR EMBL; AF201903; AAG15426.1; -;
 DR EMBL; BC013493; AAG13493.1; -;
 DR MGD; MGI:1276558; S132a2.
 DR InterPro; IPR001898; Na/sul_synth.
 DR Pfam; PF00939; Na_sulph_synth; 1.
 DR TIGRfam; TIGR00785; daas; 1.
 DR PROSITE; PS01271; NA_SULFATE; 1.
 KW Sodium transport; Symport; Transmembrane; Transport.
 FT TRANSMEM 13
 FT TRANSMEM 53 73
 FT TRANSMEM 86 106
 FT TRANSMEM 215 235
 FT TRANSMEM 264 284
 FT TRANSMEM 319 339
 FT TRANSMEM 366 386
 FT TRANSMEM 407 427
 FT TRANSMEM 445 465
 FT TRANSMEM 478 498
 FT TRANSMEM 506 526
 FT TRANSMEM 535 555
 SQ SEQUENCE 586 AA; 64110 MW; 56EAE21532833EE CRC64;
 Query Match 34.5%; Score 1033.5; DB 1; Length 586;
 Best Local Similarity 38.1%; Pred. No. 3.7e-69;
 Matches 218; Conservative 128; Mismatches 189; Indels 37; Gaps 9;
 QY 26 LVVFLVPLLCPLVMLNBSGFRQWLLVMAIVFWTEALPLYVTSMIPVAPFIMGMS 85
 DB 15 LIVLCPLPLPLPLVQTKAYCAYSILMALMLCTEALPLAVTALPIILFPLMGME 74
 QY 86 SDQCRLYFKDTLVFMGGIMVALAVEYCNHKLRLRVIQVGCSPRLHFGLLMTVMF 145
 DB 75 ASKVCLEYFKDTNLFVGLMVALAVEHNLHKLALGVLLIIGVPPALLLLGFMVLTAF 134
 QY 146 LSMISNAACTAMCPIIQAVLELQA--QGVCKINHEPQYQIVGCKNNKNE----- 195
 DB 135 LSMISNTATTAMMLPIGYALEQLQSQKDVESGNSNPSPELQEPASQKBEETKLDNGOA 194
 QY 196 ----DEP-----PYPTKITCYLGIAYASLSGCGTIIIGTATNLTKGIEARPK 242
 DB 195 VVSSEPPRAQTKBHRFSGLSLC-----ICYSASIGGIATLTGTPNLVQSQVNSIPP 250
 QY 243 NSTEQMDPTFMFVSFMSLVYTLTLTFVLOHFWGLWRPKSKEAQVQREGADVAKK 302
 DB 251 ENSNVNFAFWGFAFTWILLALLMLQLVGLVNFNFRNFGGEGEERKQA--AFQ 308
 QY 303 VIDQRYKDLGPMSTHIEIQVMTLFIQVVMYFTRKPGIFLGNAD--LLNSKD---IRNSMP 357
 DB 309 VIKQTHLLGPMSPAERAVTELFVLLVVLWFTREPPFGFGMGDTAFANKKQSMVSDGTV 368
 QY 358 TIFVVMCFMLPANYAFELRYCTRGGPVPTGPTSLITWKFIQTQVPLVLLGGGFL 417
 DB 369 AIFSLNFIIPSKPGLTEDPKPKGLKA--PPAILTKVTNDKMPNMLILLGGGFL 426
 QY 418 AEGSKQSMAGLIGNALIGLVPLNSVLLLVILVAVELTAFSSNVAITNIIPVLAEMS 477
 DB 427 AKGEESGLSKWLGDKLTPLOHVPSPATVLLSLLVAIFTECTSNVATTPLPLASMA 486
 QY 478 LAIEIHLPLVLLPAGLACSMFHLVSTPPNALVAGYANIRTKDMAIAGIGPTITITL 537
 DB 487 QAICLHPLYVMLPCTLAASLAFMLPVATPPNAIVSFGLKVKVSDMARAGFLNIIIGVLAI 546
 QY 538 FVFCOTWGLVVPNLNRPPEMAQIYAAALGN 569
 DB 547 TLSINSWSIPIF-KLDTFPTTWAYSNSTQCLLN 577

RESULT 10
 Q6AZR9
 ID Q6AZR9
 AC Q6AZR9;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 PRT; 586 AA.
 PRELIMINARY;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)

DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE S132a2-prov protein.
GN Name=s132a2-prov;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RP TISSUE=Kidney;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RL Dev. Dyn. 225:384-391(2002).
RN [2]
RP SEQUENCE FROM N.A.
RP TISSUE=Kidney;
RX PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haiech F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE FROM N.A.
RP TISSUE=Kidney;
RA Klein S., Gerhard D.S.; to the EMBL/GenBank/DBJ databases.
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC077435; AAH77435.1; -
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006814; P:sodium ion transport; IEA.
DR InterPro; IPR001898; Na/sul_sympot.
DR Pfam; PF00939; Na_sulph_symp; 1.
DR PROSITE; PS01271; NA_SULFATE; 1.
SQ SEQUENCE 586 AA; 64959 MW; 2E3646A1D81F31DB CRC64;

Query Match 34.38; Score 1028; DB 2; Length 586;
Best Local Similarity 37.18; Pred. No. 9.5e-69;
Matches 221; Conservative 121; Mismatches 187; Indels 66; Gaps 11;

QY 15 CSNPFANHKGVLVFLVPLLLPVLMLNEGAFCRCMYLLLVMAIFVWTEALPLYVTSMIP 74
DB 4 CMKWLNCKRFLILFWPFLLLPLLLPLLIQTKEASCAVVIIMAVYWCTEVPLAVTALLP 63

QY 75 IVAPPIGIMSSDQTCRLYKFDLTVMFGIMVALAVEYNLHKRLALRVIQVGCSPRR 134
DB 64 VLLFPFMEGILESCKVCQYLKDTNMLFVGLIVAVAEQWNLHKRIALKVLLIVGVRPAL 123

QY 135 LHFGILVWTFSLNWIENAACTAMWCPIIOAVLEEL----- 170
DB 124 LMLGFMGTAFSLNWIENATATTAMVFIQVRLSQLTASEDPSMLESQNTNPAALELE 183

QY 171 -----QAQGVCKINHEPQVQIVGKNKKNDEBPPYPTK-ITLCYYVLGIAYASSLGGCG 222

Db 184 TKNAIPMLQVQTVSNHGHVEVP-----EEVPEKDERKHTSGMMLC-----VCYAAISGGTA 235
QY 223 TIIGTATNLTFKGIYEAFKFNKSTEQMDPFTFMFYSVPSMLVYTLTFVFLQWFMGLWRP 282
Db 236 TLGTGPNVLVKGQFTQIFPNNGDILNFSWFGFAFNNVMIMCLSWFWLQFSGFNFK 295
QY 283 KS-----KEAQEVORREGADVAKVVDIQRYKDLGPMISHEIQVMILFIFVNVVYFRKPG 338
Db 296 KTWGCGATASEKER-----AAYSVIREYRKLGPISYAESSVLFLILLVLLWFTDFPG 349
QY 339 IFLGWADLLNSKD-----INSMPTIFVVMVCMFLPANYAFRLYCTRRGG-----PV 385
Db 350 FVVGWATILFNKDNIEYATDAVAVFAVFLFLPATKPKFGCKKARNSEDLDFEEOQ 409
QY 386 PTGPTPSLITWKFIQKVPWGLVFLGGGFALAGSGKQSGMAKLIGNALIGLKLPLN-SV 444
Db 410 DTFFSAPLLTWKVVQKQMPWSIVLLGGGFALAKGSDASGLSHWLGQQTPLHSIPWAI 459
QY 445 LLLVVLVAVFLTAFSSNVAIANIIPVLAEMSLAIEIHPVLILPAGLACSMAFHLPVS 504
Db 470 AILSLMIAVF-TECASNVATATLPLILASMSRSIEVNPILIMICTLSTSFAPMLPA 528
QY 505 TPNALVAGYANIRTKDMAIGIPPIITITLFCVTCQWGLVVPVNLNSFPPEWA 559
Db 529 TPNNAIVFSYGHRLVSDVMVKTGIVMNIIGICTTISINTWGRPMF-SLDTFPQMA 582

RESULT 11
S132_HUMAN STANDARD; PRT; 592 AA.
ID S132_HUMAN
AC Q13183;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Solute carrier family 13, member 2 (Renal sodium/dicarboxylate
DE cotransporter) (Na(+)/dicarboxylate cotransporter 1) (NaDC-1).
GN Name=SLC13A2; Synonyms=NADC1, SDCT1;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=96199379; PubMed=8967342;
RA Fajor A.M.;
RT "Molecular cloning and functional expression of a sodium-dicarboxylate
RT cotransporter from human kidney.";
RL Am. J. Physiol. 270:F642-F648(1996).
CC -I- FUNCTION: Cotransport of sodium ions and dicarboxylates such as
CC succinate and citrate.
CC -I- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -I- SIMILARITY: Belongs to the SLC13A transporter family. NADC
CC subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U26209; AAA98504.1; -;
DR Genew; HGNC:10917; SLC13A2.
DR MIM; 604148; -;
DR GO; GO:0005887; C:integral to plasma membrane; TAS.
DR GO; GO:0005624; C:membrane fraction; TAS.
DR GO; GO:0015361; F:low affinity sodium:dicarboxylate symporter. . .; TAS.
DR GO; GO:0006810; P:transport; TAS.
DR InterPro; IPR001898; Na/sul_sympot.
DR Pfam; PF00939; Na_sulph_symp; 1.
DR

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2].
RP SEQUENCE FROM N.A.
RC STRAIN=AB; TISSUE=Whole body;
RA Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC044437; AAH44437.1; -
DR ZFIN; ZDB-GENE-040426-2803; zgc:55601.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006814; P:sodium ion transport; IEA.
DR InterPro; IPR001898; Na/sul_synth.
DR Pfam; PF00939; Na_sulph_synth; 1.
SQ SEQUENCE 613 AA; 67594 MW; F6B73FE343F288C8 CRC64;
Query Match 34.2%; Score 1023.5; DB 2; Length 613;
Best Local Similarity 36.5%; Pred. No. 2.2e-68;
Matches 217; Conservative 122; Mismatches 192; Indels 63; Gaps 9;
QY 22 HWKGLVPLVPLLCPLPMLNLEGAFCRCMYLLVNAIFWTEALPLYVTSMIPVAPPIM 81
DB 16 HRCNLCILCVPLILPLPLVPTPEARGCGFAILLMALEFWCTECMPLAITALLPVVLPFM 75
QY 82 GIMSSDQTCRLYFKDTLVFMFGIMVALAVEYCNLHKRLALRVIOVGCSPRLHFLGIM 141
DB 76 GIMSEGEVCVQLKDTNMLFVGLLVAIVAEHNLHRIALSVLLVGVRRPALLMLGPMI 135
QY 142 VTMLSMWISNAACATMCPITIOAVLEEL-----QAQGVCKINHEBPQ 183
DB 136 VTAFLSMWSISNTATTAMLPISQAVLEQLSATEADSEKELREGQVNAQFELTEVNIKQP 195
QY 184 YQIVGKNKNE-----DEPPYTKITLCYLGIAVASSLGGCGTIG 226
DB 196 LDNTQGEKPNNAADVEGINALSERRRKAREAKY-LRLFKGMSLSVCYSASIGGTATLTG 254
QY 227 TATNLTFGIYEAPKFNSTEQMDPTFMFVSVPMSLVTLTTFVLOWHFMGLWRPKSKE 286
DB 255 TTPNLILKGQDEIFPDNDNVINFASWFGFAPPNVLMVLVSLWLCQCYLGF---NFKQ 311
QY 287 AQEVQREGADVAKVIDQRYKDLGPMSTHEIOVMILFIFMVMYFTRKGIPLGWADL 346
DB 312 SFGCGTKNEGDKAYKWKNEKYLKGLPMSPAGAVLVIFILVILWFTREGFMFGWATE 371
QY 347 L---NSKDIRNSMPTIFVVMCFMLPANYAFURYCT-----RRGGVPVPTG----- 388
372 LFNKNGQVVTGTVAFIMSTLFFVIPSVDL--CSIKVEQDEBAVGEQDGVKEKKR 429
399 ---PPSLITWKFQTKVPGVFLVLLGGFALAEKSGKAGKILGNALIGKVLPSVL 445
430 KLKGTPTLLNKKVHREPMVNLVLLGGFALAEKSGESGLVSWLQSLPSLPPPAI 489
446 LLVTLVAVFLTARSSNAIANIIPVLAESLATEIHPVLILPAGLACSNAPFLPVST 505
490 SIILCLLVGTTECSNTATTLFPLILASHMATTIGLHPLVLMCTTISASLAFMLPVAT 549
506 PPNALVAGYANIRTKDMAIGPTIITITLTFVFCQTWGLVVPVNLNSFPPEWA 559
550 PPNALFAPSYGNLKVLDMAKAGFILNIIGILCINLGINTWGMAMF-KLDTFFPSWA 602
RESULT 15
ID Q6P3L2
AC Q6P3L2; PRELIMINARY; PRT; 613 AA.
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE zgc:77607 protein.

GN ORFNames=zgc:77607;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1].
RP SEQUENCE FROM N.A.
RC TISSUE=Whole;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udgin T.B., Tohiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2].
RP SEQUENCE FROM N.A.
RC TISSUE=Whole;
RA Strausberg R.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC063947; AAH63947.1; -
DR ZFIN; ZDB-GENE-040426-2389; zgc:77607.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006814; P:sodium ion transport; IEA.
DR InterPro; IPR001898; Na/sul_synth.
DR Pfam; PF00939; Na_sulph_synth; 1.
SQ SEQUENCE 613 AA; 67653 MW; 8A4B42760D510C6B CRC64;
Query Match 34.2%; Score 1022.5; DB 2; Length 613;
Best Local Similarity 36.5%; Pred. No. 2.6e-68;
Matches 217; Conservative 122; Mismatches 192; Indels 63; Gaps 9;
QY 22 HWKGLVPLVPLLCPLPMLNLEGAFCRCMYLLVNAIFWTEALPLYVTSMIPVAPPIM 81
DB 16 HRCNLCILCVPLILPLPLVPTPEARGCGFAILLMALEFWCTECMPLAITALLPVVLPFM 75
QY 82 GIMSSDQTCRLYFKDTLVFMFGIMVALAVEYCNLHKRLALRVIOVGCSPRLHFLGIM 141
DB 76 GIMSEGEVCVQLKDTNMLFVGLLVAIVAEHNLHRIALSVLLVGVRRPALLMLGPMI 135
QY 142 VTMLSMWISNAACATMCPITIOAVLEEL-----QAQGVCKINHEBPQ 183
DB 136 VTAFLSMWSISNTATTAMLPISQAVLEQLSATEADSEKELREGQVNAQFELTEVNIKQP 195
QY 184 YQIVGKNKNE-----DEPPYTKITLCYLGIAVASSLGGCGTIG 226
DB 196 LDNTQGEKPNNAADVEGINALSERRRKAREAKY-LRLFKGMSLSVCYSASIGGTATLTG 254
QY 227 TATNLTFGIYEAPKFNSTEQMDPTFMFVSVPMSLVTLTTFVLOWHFMGLWRPKSKE 286
DB 255 TTPNLILKGQDEIFPDNDNVINFASWFGFAPPNVLMVLVSLWLCQCYLGF---NFKQ 311
QY 287 AQEVQREGADVAKVIDQRYKDLGPMSTHEIOVMILFIFMVMYFTRKGIPLGWADL 346
DB 312 SFGCGTKNEGDKAYKWKNEKYLKGLPMSPAGAVLVIFILVILWFTREGFMFGWATE 371
QY 347 L---NSKDIRNSMPTIFVVMCFMLPANYAFURYCT-----RRGGVPVPTG----- 388

```

| | : : : || : : : | | | : |
Db 372 LFNKNGQVVDGTVAIFMSTLFFVIPSVDFL--CSIKYEEQDEEAVGEQDQGEKEKKR 429
Qy 389 ---PTPSLITWKFIQTQKVPWGLYFLGGGFALAEQSKQSGMAKLI GNALIGLKVLPNSVL 445
Db 430 KLKGTPTLLNWKVHERMPWNI VLLGGGFALASGEESGLSVWLGGQSUSPLKSI PPFAT 489
Qy 446 LLVVILVAVELTAFSSNVAIANIIPVLAEMSLAIEIHPYLIIPAGLACSMAFHLPVST 505
Db 490 SIILCLLVGTFTCCSSNTATTLFLPILASMAATTIGLHPLYVMLPCTISASLAFMLPVAT 549
Qy 506 PPNALVAGYANIRTKDMAIAGIGFTIITITL FVFCOTWGLVYVYPNLNSPPEWA 559
Db 550 PPNAIAFSYGNLKYLDMAKAGFILNIIGILCINLGINTWGMAMF-KLDTPPSWA 602
```

Search completed: June 30, 2005, 08:58:41
Job time : 82.9736 secs

GenCore version 5.i.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2005, 08:28:14 ; Search time 79.2935 Seconds
(without alignments)
2741.199 Million cell updates/sec

Title: US-10-017-479A-3
Perfect score: 2897
Sequence: 1 MABPGQRKFLVGRCCIFHW.....FPESKSPDWAKEIKNQTKI 562

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : A_Geneseq_16Dec04: *
1: Geneseqp1980s: *
2: Geneseqp1990s: *
3: Geneseqp2000s: *
4: Geneseqp2001s: *
5: Geneseqp2002s: *
6: Geneseqp2003as: *
7: Geneseqp2003bs: *
8: Geneseqp2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2114.5	73.0	908	4	ABB61068 Drosophila
2	1382.5	47.7	572	4	ABB60315 Drosophila
3	1382.5	47.7	572	4	ABB66903 Drosophila
4	1382.5	47.7	572	5	ABB79611 Drosophila
5	1382.5	47.7	572	8	ADP64789 Drosophila
6	1374.5	47.4	572	6	ABR40099 drosophila
7	885	30.5	568	7	ADK51050 Human NOV
8	885	30.5	568	8	ADH42445 Human NOV
9	885	30.5	568	8	ADN61775 Human NOV
10	885	30.5	568	8	ABR57023 Human NOV
11	875	30.2	568	6	ABR57023 Human NOV
12	875	30.2	568	7	ADK51052 Human NOV
13	875	30.2	568	7	ADK51052 Human NOV
14	875	30.2	568	8	ADH42441 Human NOV
15	875	30.2	568	8	ADP64793 Human NOV
16	875	30.2	568	8	ADH42443 Human NOV
17	874.5	30.2	568	7	ADK51048 Human NOV
18	871	30.1	568	6	ABR40097 Human NOV
19	871	30.1	568	7	ADK51052 Human NOV
20	868	30.0	616	5	ABU65064 Human NOV
21	868	30.0	616	5	ABU65063 Human NOV
22	868	30.0	616	7	ADK51048 Human NOV
23	868	30.0	616	8	ADH42447 Human NOV
24	868	30.0	616	8	ADN61777 Human NOV
25	868	30.0	616	8	ADN61779 Human NOV

26	866	29.9	568	6	ABG75835	Abg75835 Transport
27	863	29.8	568	5	AAU79946	AAU79946 Human tra
28	863	29.8	568	8	ADG16993	ADG16993 Human tra
29	856.5	29.6	572	8	ADP64791	ADP64791 Rat INDY
30	852.5	29.4	619	8	ADG16995	ADG16995 African c
31	849.5	29.3	581	8	ADP64799	ADP64799 Zebrafish
32	844	29.1	592	6	ABH82950	ABH82950 Human SLC
33	844	29.1	592	6	ABH82951	ABH82951 Human SLC
34	844	29.1	592	6	ABR40100	ABR40100 Human Ren
35	844	29.1	592	6	ABR57025	ABR57025 Human NaD
36	844	29.1	592	7	ADP63755	ADP63755 Human Pro
37	836.5	28.9	539	5	AAE21181	AAE21181 Human TRI
38	833.5	28.8	531	6	ABG75826	ABG75826 Transport
39	831.5	28.7	602	6	ABH82952	ABH82952 Human SLC
40	831.5	28.7	602	6	ABR40101	ABR40101 Human Ren
41	831.5	28.7	602	6	ABG75837	ABG75837 Transport
42	814.5	28.1	572	6	ABR57024	ABR57024 Mouse TCH
43	814.5	28.1	572	8	ADP64797	ADP64797 Mouse Na+
44	812.5	28.0	551	8	ADR09550	ADR09550 Human pro
45	743	25.6	626	3	AAH36167	AAH36167 Novel hum
46	742.5	25.6	627	3	AAH36161	AAH36161 Novel hum
47	742	25.6	522	5	ABU65065	ABU65065 Human NOV
48	742	25.6	522	7	ADK51054	ADK51054 Human NOV
49	742	25.6	522	8	ADH42449	ADH42449 Novel hum
50	742	25.6	522	8	ADN61781	ADN61781 Human nov
51	741	25.6	626	3	AAH36164	AAH36164 Novel hum
52	740.5	25.6	627	3	AAH23625	AAH23625 Human sec
53	740.5	25.6	627	3	AAH36158	AAH36158 Novel hum
54	738.5	25.5	627	3	AAH42213	AAH42213 Human ORF
55	738.5	25.5	627	5	ABH97450	ABH97450 Novel hum
56	737	25.4	557	8	ADP29692	ADP29692 Human sec
57	737	25.4	626	3	AAH36168	AAH36168 Novel hum
58	736.5	25.4	627	3	AAH36162	AAH36162 Novel hum
59	735	25.4	626	3	AAH36165	AAH36165 Novel hum
60	734.5	25.4	627	3	AAH36159	AAH36159 Novel hum
61	730.5	25.2	519	8	ABM84184	ABM84184 Human tra
62	721.5	24.9	519	8	ADH22543	ADH22543 Human tra
63	708	24.4	533	8	ABM84303	ABM84303 Human dia
64	698	24.1	520	7	ADM05177	ADM05177 Human pro
65	696	24.0	595	5	AAE22910	AAE22910 Human tra
66	696	24.0	595	5	AAO21807	AAO21807 Lung-spec
67	696	24.0	595	6	ABH82949	ABH82949 Human SLC
68	696	24.0	595	7	AD121045	AD121045 Novel hum
69	696	24.0	595	8	ADO78130	ADO78130 Human SLC
70	693.5	23.9	516	5	ABU65066	ABU65066 Human NOV
71	693.5	23.9	516	7	ADK51056	ADK51056 Human NOV
72	693.5	23.9	516	8	ADH42451	ADH42451 Novel hum
73	679	23.4	477	8	ADN61783	ADN61783 Human nov
74	671	23.2	551	8	ADP64795	ADP64795 Nematode
75	624	21.5	580	3	AAH36169	AAH36169 Novel hum
76	623.5	21.5	581	3	AAH36163	AAH36163 Novel hum
77	622	21.5	580	3	AAH36166	AAH36166 Novel hum
78	621.5	21.5	581	3	AAH36160	AAH36160 Novel hum
79	601.5	20.8	552	6	AAO31005	AAO31005 Human tra
80	588	20.3	543	6	ADA48682	ADA48682 Rice prot
81	588	20.3	543	7	ADC08241	ADC08241 Rice prot
82	551	19.0	510	4	AAG90000	AAG90000 C glutami
83	549	19.0	432	6	ABM71452	ABM71452 Staphyloc
84	545	18.8	502	6	ADA48226	ADA48226 Rice prot
85	545	18.8	552	2	AAW98815	AAW98815 H. pylori
86	544	18.8	527	4	AAH76767	AAH76767 Corynebact
87	544	18.8	540	3	AAH51055	AAH51055 Arabidops
88	544	18.8	557	3	AAH51054	AAH51054 Arabidops
89	543	18.7	540	3	AAH51053	AAH51053 Arabidops
90	543	18.7	558	3	AAH51056	AAH51056 Arabidops
91	539.5	18.6	466	3	AAH51057	AAH51057 Arabidops
92	538.5	18.6	466	3	AAH51058	AAH51058 Arabidops
93	535	18.5	407	5	AAU91115	AAU91115 Human sec
94	535	18.5	407	5	ABG65198	ABG65198 Human alb
95	535	18.5	407	5	ABG65199	ABG65199 Human alb
96	511.5	17.7	377	5	AAU91090	AAU91090 Human sec
97	511.5	17.7	377	5	ABG65200	ABG65200 Human alb
98	511.5	17.7	377	8	ADL78467	ADL78467 Albumin f

99 487.5 16.8 382 7 ADE09041 Novel pro

100 459.5 15.9 462 6 ABM70253 Photorhab

ALIGNMENTS

RESULT 1

ABBB61068

ID ABB61068 standard; protein; 908 AA.

XX ABB61068;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 9996.

DE Drosophila melanogaster polypeptide; cell signalling; insecticide;

KW pharmaceutical.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

PR 11-JUL-2000; 2000US-00614150.

XX (PEKE) PE CORP NY.

FA Venter JC, Adams M, Li PWD, Myers EW;

PI WPI; 2001-656860/75.

DR N-PSDB; ABL05171.

XX New isolated nucleic acid detection reagent for detecting 1000 or more

PT genes from Drosophila and for elucidating cell signalling and cell-cell

PT interactions.

XX Disclosure; SEQ ID NO 9996; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent

CC capable of detecting 1000 or more genes from Drosophila. The invention is

CC useful in developmental biology and in elucidating cell signalling and

CC cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA

CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-

CC ABB72072). The sequence data for this patent did not form part of the

CC printed specification, but was obtained in electronic format directly

CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 908 AA;

XX Query Match 73.0%; Score 2114.5; DB 4; Length 908;

XX Best Local Similarity 49.8%; Pred. No. 2.1e-210;

XX Matches 452; Conservative 20; Mismatches 38; Indels 397; Gaps 2;

QY 44 MAEFKCLWLVITMALLWITETLPIYVTVLFPVFCPLGLVNASIVCKQYFTDITVIVFLG 103

DB 1 MAEFKCLWLVITMALLWITETLPIYVTVLFPVFCPLGLVNASIVCKQYFTDITVIVFLG 60

QY 104 GLIIVALGIEYNLHTRIALRVIRIVGGSPRRLFVGLMSVSTFMGLWISNSAGTAMMCPIV 163

DB 61 GLIIVALGIEYNLHTRIALRVIRIVGGSPRRLFVGLMSVSTFMGLWISNSAGTAMMCPIV 120

QY 164 KALVNELDTNKIFPVYMTQEEBPEEPPHPSKITVAFYAGIAYASSIGGLTIGTGT 223

DB 121 KALVNELDTNKIFPVYMTQEEBPEEPPHPSKITVAFYAGIAYASSIGGLTIGTGT 180

QY 224 NLVPRGIYTERPPTSTVEITFANFMFYSIPLMVI VNVTLVIIAFILITHMGLFRPNRSKTGK 283

DB |||||

QY 181 NLVPRGIYTERPPTSTVEITFANFMFYSIPLMVI VNVTLVIIAFILITHMGLFRPNRSKTGK 240

DB |||||

QY 284 IIAEANTNRKLMEDVLRQRHIDLGPMSCHEIQMAIAFAPMIVLLITRKPFGFVPGWSDLIN 343

DB |||||

QY 241 IIAEANTNRKLMEDVLRQRHIDLGPMSCHEIQMAIAFAPMIVLLITRKPFGFVPGWSDLIN 300

DB |||||

QY 344 RKVGSASGLSFIVLLIFALPTQYTFYFKYCCGKGPFTAQIDAIIISWEVLRNIPWGLLF 403

DB |||||

QY 301 RKVGSASGLSFIVLLIFALPTQYTFYFKYCCGKGPFTAQIDAIIISWEVLRNIPWGLLF 360

DB |||||

QY 404 LGGGFALAVASRETGLNIMISKAMQVLLIGLNPVIVQSIT 443

DB |||||

QY 361 LGGGFALAVASRETGLNIMISKAMQVLLIGLNPVIVQSITADGKIGGGQAVKRCGCFHWR 420

DB |||||

QY 444 443

DB 421 GKATAMIPLITLPIIMYGVLENLNASYGVVDILESTSEKVCSEYFSDTVVMFIGLLIAL 480

QY 444 443

DB 481 AIEYNLHORIAALNTILIVGCSPRRLHFGLVNMTCFISLWISNSAATAMCPIVKAVLNE 540

QY 444 443

DB 541 METQNIIFAIYKTOBEEPEVEEGDPHPSTISMAFYGCIAYSSIGGCGTLTGCTNLTGK 600

QY 444 443

DB 601 LYDTRPPNSDEKIDPFIYMAYSPFVVLVILFTVPSLQVTHMGLFRPNRSKIGQEVKGA 660

QY 444 443

DB 661 ESQDVVQVIVKORKAELGPMSCHEIQVGLL FVLMIFLLETRKPGPPGWADFLNKAIGS 720

QY 444 448

DB 721 GPPVFPATILLFALPTQYTFYFKYCCGKAPPGQTLDACLSGWGFALAEGRSVGMAKMLGE 780

QY 449 486

DB 781 SLAFAGEMHSVLVISMCIISLFCFASNAVAICNLIPIFSEMAIAIEVHPMKLTFPPAA 840

QY 487 LGISMYVFLPVSTPPNAIVTOVAHIKTYFACCGIVPTIIGISVALVNTNMTWGLIIFPES 546

DB 841 LACSLAFHLFPVSTPPNAILISGFTGLTKYMAIAGILPTCWAFLCLLFTGTGVTMTLIVPGT 900

QY 547 KSFPDWA 553

DB 901 TEPFSA 907

RESULT 2

ABBB60315

ID ABB60315 standard; protein; 572 AA.

XX ABB60315;

XX 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 7737.

KW Drosophila; developmental biology; cell signalling; insecticide;

XX pharmaceutical.

OS Drosophila melanogaster.

XX WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US009231.

DB 240 RFKNSTQMDPTFMFVSFMSLV--YTLTFLVFLQWFMGLWRPKSKEAQEVQRGREGA 297
 QY 293 KLMDVLRQRHIDLGPMSCHEIQMAIAFAFMIVLLITRKPGFVPGWSDLINRKVGSAGS 352
 DB 298 DVAKKVIDQYKDLGPMSCHEIQMAIAFAFMIVLLITRKPGFVPGWSDLINRKVGSAGS 357
 QY 353 LSFIVLLIFALPTQYTFKVCCKG-GPPTAQADAILSWYVLRNIPWGLLFLGGGFAL 411
 DB 358 TIFVVMCFMPLPANYAFRLYCTRGGVPPTGPTSLITWKFIQTKVPWGLVFLGGGFAL 417
 QY 412 AVASRETLNIMISKAMQVLIGLNPVQSIITFVLNFPFSAFNANVVANIVLPILCEMS 471
 DB 418 AEGSKQSGMAKLGNALIGLKVLPNSVLLVILVAVFLTAFSSNAIANIIPVLAEMS 477
 QY 472 LALEHPLILTLPACLGISMVYFLPSTPPNAIVTOYAHIKTKYFACCGIVPTIIGISVA 531
 DB 478 LAIEIHPVLILPAGLACSMFHLPVSTPPNALVAGYANIRTKDMAIAGIGPTIITIIL 537
 QY 532 LVNTNTWGLIIFPESKSFDPWAK 554
 DB 538 FVFCQTWGLVVPNLNSFPPEWAQ 560

RESULT 4
 ABB79611
 ID ABB79611 standard; protein; 572 AA.
 XX
 AC ABB79611;
 XX
 DT 21-OCT-2002 (first entry)
 XX
 DE Drosophila INDY protein dicarboxylate transporter.
 XX
 KW Indy; dicarboxylate transporter; life span; longevity; obesity;
 KW anorectic; caloric restriction; transgenic animal; gene therapy.
 XX
 OS Drosophila melanogaster.
 XX
 FN WO200259310-A2.
 XX
 PD 01-AUG-2002.
 XX
 PF 12-DEC-2001; 2001WO-US048130.
 XX
 PR 12-DEC-2000; 2000US-0255013P.
 XX
 PA (UYCO-) UNIV CONNECTICUT.
 XX
 PI Regina B, Reenan RA, Helfand SL;
 XX
 DR WPI; 2002-599787/64.
 XX
 DR N-PSDB; ABB84439.
 XX

Indy polynucleotide, useful for diagnosing or treating body weight disorders, e.g. obesity, metabolic maintenance disorders, or the symptoms of aging to extend the life span of an organism.

Claim 14; Page 81-83; 83pp; English.

The present sequence is the protein sequence of the Drosophila melanogaster INDY protein, which is encoded by the Indy gene (see ABB84439) involved in increased life span. The protein has similarity to dicarboxylate transporters such as those from human and rat, and represents a new class of dicarboxylate transporters that are not inhibited by phloretin. Identification of the Indy gene resulted from the observation that particular mutations in the gene caused an increase in the life span of the fly carrying the mutation. As a result of this finding, it is now possible to identify and/or isolate Drosophila lines with longer life spans, as well as to identify agents that contribute to longer life span. It is also possible to isolate genes involved in, and which have an effect on longevity, as well as proteins encoded by these genes. The invention provides Indy polynucleotides, proteins, anti-INDY

CC antibodies, antagonists that inhibit Indy activity or expression and
 CC agonists that increase Indy activity or expression, and their use in the
 CC diagnosis or treatment of body weight disorders, such as obesity and
 CC metabolic maintenance disorders, or longevity in humans and animals.
 CC Antagonists include at least a portion of the Indy gene sequence, an
 CC antisense oligonucleotide, a ribozyme, a triple helix-forming molecule, a
 CC double-stranded interfering RNA, an anti-Indy antibody, or a mixture of
 CC these. Methods of calorically restricting an organism and of extending
 CC the lifespan of an organism by administering the antagonist are claimed
 XX
 SQ Sequence 572 AA;

Query Match 47.7%; Score 1382.5; DB 5; Length 572;
 Best Local Similarity 46.5%; Pred. No. 1,9e-134;
 Matches 262; Conservative 120; Mismatches 164; Indels 17; Gaps 7;
 QY 3 EFGQRKFLVGRCCIF---HWRGKASIIIPILITLILYGFQTDMAEFKCLMIVTMALL 59
 DB 4 EIGEOPQPPV-KCSNFFPANHKGVLVFLVCLLPVLL---NEGABFCMVLVLLVNAIF 59
 QY 60 WITETLPIVVTALFPLVFCPLGLGNASIVCKQYFTDTIIVFLGLLIVAGIEYSNLHTR 119
 DB 60 WYTEALPLYVTSMIPIVAPPIMGINSDDQTCRLYFKDILVPMFGGIMVALAVEYCNLHKR 119
 QY 120 IALRVIRIVGSPRRLFVGLMSVTFMGLWISNAGTAMMCFIVKALVNELDTNKIPVY 179
 DB 120 LALRVIIQVGCSPRRLHFLGIMVTFMSWISNAACTAMMCPILQAVLBELOAQGVCKIN 179
 QY 180 MTQEEPV-----EKGEPHPSKITVAFYAGYASSIGGLTGLTGTNLVFRGIYTE 233
 DB 180 HEFYQIVGNGKKNNEDEPPYPTKITLCYVLGAYASSIGGCGTIIIGTATNLTFKGIYA 239
 QY 234 RPPTSVETTFANMFYSIPLMVIVNVLIVIAFLITH-MGLFRPNSTKTKIIAEANTNR 292
 DB 240 RFKNSTQMDPTFMFVSFMSLV--YTLTFLVFLQWFMGLWRPKSKEAQEVQRGREGA 297
 QY 293 KLMDVLRQRHIDLGPMSCHEIQMAIAFAFMIVLLITRKPGFVPGWSDLINRKVGSAGS 352
 DB 298 DVAKKVIDQYKDLGPMSCHEIQMAIAFAFMIVLLITRKPGFVPGWSDLINRKVGSAGS 357
 QY 353 LSFIVLLIFALPTQYTFKVCCKG-GPPTAQADAILSWYVLRNIPWGLLFLGGGFAL 411
 DB 358 TIFVVMCFMPLPANYAFRLYCTRGGVPPTGPTSLITWKFIQTKVPWGLVFLGGGFAL 417
 QY 412 AVASRETLNIMISKAMQVLIGLNPVQSIITFVLNFPFSAFNANVVANIVLPILCEMS 471
 DB 418 AEGSKQSGMAKLGNALIGLKVLPNSVLLVILVAVFLTAFSSNAIANIIPVLAEMS 477
 QY 472 LALEHPLILTLPACLGISMVYFLPSTPPNAIVTOYAHIKTKYFACCGIVPTIIGISVA 531
 DB 478 LAIEIHPVLILPAGLACSMFHLPVSTPPNALVAGYANIRTKDMAIAGIGPTIITIIL 537
 QY 532 LVNTNTWGLIIFPESKSFDPWAK 554
 DB 538 FVFCQTWGLVVPNLNSFPPEWAQ 560

RESULT 5

ADP64789
 ID ADP64789 standard; protein; 572 AA.

XX

AC ADP64789;

XX

DT 26-AUG-2004 (first entry)

XX Drosophila INDY, an Na⁺-coupled citrate transporter protein.

XX sodium-coupled citrate transporter; transmembrane citrate transporter;
 KW lifespan; weight reduction; weight gain prevention; blood cholesterol;
 KW triglyceride; low density lipopolsaccharide; glucose; obesity;
 KW hyperlipidemia; hypercholesterolemia; INDY protein.

XX Drosophila melanogaster.

OS

Db 180 HEPOYQIVGKKNKNEDEPPYPTKITLCYVLGIAYASSLGGCGTIIIGTATNLTFRKGIYEA 239
Qy 234 RFTSTVEITFANPMFYSIDLMVIVNVLIIAFLITH-MGLPRNSKTKGIITAEANTNR 292
Db 240 RFKNSTEQMDPFTFMFYSVPSMLV--YTLTTFVFLQWHFGLWRPKSKEAQEVQREGA 297
Qy 293 KLMEVLQRHIDLPGMSCHIEIQMAIAFAFMIVILLITRKPGFVPGMSDLINRKVGSASG 352
Db 298 DVAKVIDQYKDLGPNISHEIQMILFIWVMYFTRKEGIFGLWADLLNSKDINSMP 357
Qy 353 LSFIVLIFALPQYTFYFKYCCGK-GPFTAQADAILSWBYLNRNIPWGLFLFLGGGFAL 411
Db 358 TIFVVMCFMLPANYAFRLYCTRGRGVPPTPSLTITWKFIQTKVPWGLVFLGGGFAL 417
Qy 412 AVASRETLNIMISKAMQVLIGLIPNVVQSITFVLANFFSAFNANVVANVIVLPILCEMS 471
Db 418 AEGSKQSGMAKLIGNALIGLKVLPNSVLLLVILVAFLVAFSSNVAIANIIPVLAEMS 477
Qy 472 LALELHPLILTLPACLGISMVYFLPVSTPPNAIVTQVAHIKTKYFACCGIVPTIIIGISVA 531
Db 478 LAIEIHPYLLLPAGLACSAFHLPVSTPPNALVAGYANIRTKDMAIAGIPTIIITL 537
Qy 532 LVNTNTWGLIIPESKSPDPAK 554
Db 538 FVFCQTWGLVVPNLNFSPEWAQ 560

RESULT 7

ABU65062
ID ABU65062 standard; protein; 568 AA.

AC ABU65062;

DT 20-MAY-2003 (first entry)

XX Human NOV14a protein.

XX NOVX; cytostatic; cardiant; antiarteriosclerotic; antiasthmatic; cancer;
KW hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;
KW human.

OS Homo sapiens.

PN WO200272757-A2.

XX 19-SEP-2002.

XX 08-MAR-2002; 2002WO-US006908.

XX 08-MAR-2001; 2001US-0274101P.

PR 08-MAR-2001; 2001US-0274194P.

PR 08-MAR-2001; 2001US-0274281P.

PR 08-MAR-2001; 2001US-0274322P.

PR 09-MAR-2001; 2001US-0274849P.

PR 12-MAR-2001; 2001US-0275235P.

PR 13-MAR-2001; 2001US-0275578P.

PR 13-MAR-2001; 2001US-0275601P.

PR 14-MAR-2001; 2001US-0276000P.

PR 16-MAR-2001; 2001US-0276776P.

PR 19-MAR-2001; 2001US-0276994P.

PR 20-MAR-2001; 2001US-0277239P.

PR 20-MAR-2001; 2001US-0277321P.

PR 21-MAR-2001; 2001US-0277791P.

PR 22-MAR-2001; 2001US-0277833P.

PR 23-MAR-2001; 2001US-0278152P.

PR 26-MAR-2001; 2001US-0278894P.

PR 27-MAR-2001; 2001US-0278999P.

PR 27-MAR-2001; 2001US-0279036P.

PR 28-MAR-2001; 2001US-0279344P.

PR 30-MAR-2001; 2001US-0277338P.

PR 30-MAR-2001; 2001US-0279995P.
PR 30-MAR-2001; 2001US-0280233P.
PR 02-APR-2001; 2001US-0280802P.
PR 02-APR-2001; 2001US-0280822P.
PR 02-APR-2001; 2001US-0280900P.
PR 04-APR-2001; 2001US-0281194P.
PR 13-APR-2001; 2001US-0283675P.
PR 30-APR-2001; 2001US-0287424P.
PR 02-MAY-2001; 2001US-0288066P.
PR 03-MAY-2001; 2001US-0288342P.
PR 03-MAY-2001; 2001US-0288528P.
PR 15-MAY-2001; 2001US-0291190P.
PR 16-MAY-2001; 2001US-0291099P.
PR 16-MAY-2001; 2001US-0291240P.
PR 30-MAY-2001; 2001US-0294485P.
PR 31-MAY-2001; 2001US-0294889P.
PR 31-MAY-2001; 2001US-0294899P.
PR 18-JUN-2001; 2001US-0299027P.
PR 19-JUN-2001; 2001US-0299303P.
PR 19-JUN-2001; 2001US-0299310P.
PR 10-JUL-2001; 2001US-0304354P.
PR 31-JUL-2001; 2001US-0309198P.
PR 16-AUG-2001; 2001US-0312903P.
PR 10-SEP-2001; 2001US-0318462P.
PR 12-SEP-2001; 2001US-0318770P.
PR 27-SEP-2001; 2001US-0325430P.
PR 27-SEP-2001; 2001US-0325681P.
PR 18-OCT-2001; 2001US-0330380P.
PR 31-OCT-2001; 2001US-0335301P.
PR 14-NOV-2001; 2001US-0332172P.
PR 14-NOV-2001; 2001US-033271P.
PR 14-NOV-2001; 2001US-033272P.
PR 14-NOV-2001; 2001US-0333184P.
PR 21-NOV-2001; 2001US-0333272P.
PR 21-NOV-2001; 2001US-0332094P.
PR 03-DEC-2001; 2001US-0337426P.
PR 03-DEC-2001; 2001US-0338092P.
PR 04-DEC-2001; 2001US-0337185P.
PR 03-JAN-2002; 2001US-0345705P.
PR 07-MAR-2002; 2002US-00092900.
XX (CURA-) CURAGEN CORP.
XX Padigar M, Spytek KA, Shenoy SG, Taupier RJ, Pena CEA, Li L;
XX Zerhusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;
PI Patturajan M, Gangolli E, Vernet CAM, Guo X, Tchernev V;
PI Fernandes ER, Casman SJ, Malyankar UM, Gerlach V, Liu Y, Anderson D;
PI Spaderna SK, Catterton E, Burgess C, Leite M, Zhong H, Alsobrook JP;
PI Lepley DM, Rieger DK;
XX WPI; 2002-723332/78.
DR N-PSDB; ABX97029.
XX NOVX polypeptides and polynucleotides, useful for preventing or treating
PT a disorder associated with aberrant NOVX expression or activity e.g.,
PT cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
PT asthma.
XX Claim 1; Page 133; 1103pp; English.
PS This invention describes novel human NOVX polypeptides which have
XX cytostatic, cardiant, antiarteriosclerotic, antiasthmatic and hypotensive
CC activity. Pharmaceutical compositions comprising the NOVX proteins or
CC nucleic acid molecules or NOVX antibodies are useful for preventing or
CC treating a disorder associated with aberrant NOVX expression or activity
CC e.g. cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
CC asthma. The products of the invention can be used for gene therapy or in
CC a vaccine. ABU65041-ABU65218 represent the NOVX polypeptides encoded by
CC ABX97008-ABX97185

SQ Sequence 568 AA;

Query Match

30.5%; Score 885; DB 5; Length 568;

Best Local Similarity 35.6%; Pred. No. 1.2e-82;		Matches 197; Conservative 101; Mismatches 219; Indels 36; Gaps 9;	
QY	27	IIPLTLPILYIGFQDMAEFKLMIVTMALLWITETLPYVVTALFPLVFCPLGLVNA	86
Db	19	VTPLLLPLVI-----LMPAKVSCAVIILMAIYVCTEVIPLAVTSLMPVLLFPLQILDS	74
QY	87	SVVCKQYFDTIVVFLGGLIVALGIEYSNLHRIALRVIRVGGSPRRFLVGLMSVSTFM	146
Db	75	ROVCVQYKMDTNMLFLGGLIIVAVAVERNLHRIALRTLLMWGAKPARMLGFMGVFTALL	134
QY	147	GLWISNAGTAMCPIVKNVNELDN-----KIPFVYMTQEEBP-VEE	189
Db	135	SMWISNTATTAMVPIVEAILQOMEATSAATAGLELVDKGAKELPGSQVIFEGPTLQ	194
QY	190	GEPPHPSKITVAFYAGIAYASSIGGLTIGTGNLFRGIYTERPPTSTVETITFANFMF	249
Db	195	QEDQERKRLCKAWTLCICVYASIGGTATLTGTGNVLLGQNNELFPDSKDLVNFASWFA	254
QY	250	YSIPLMIVNVTLVIIAFL-ITHMGLFRPNSKTGKIIAEANTNRKLMEDVLRQRHIDLGP	308
Db	255	FAPFNNLV-----MLLFAWLQFVYMFSSFKKSGCGLESKKNKAALKVQEEYRKLGP	310
QY	309	MSCHIQMAIAFAFMIVLLITRKPGFVCGWSDLI-----NRKVGSASGLSFIVLILFALP	364
Db	311	LSFABINVLICFPLLVILWFSRDPGFMGWLTVAVVEGETKYVSDATVAIFVATLLFIVP	370
QY	365	TOYTFKVC-----GKGPFTAQDAIDALSWEYVURNPWGLLFLGGFALAVASREGL	420
Db	371	SQKPFNFRSQTEEGKSPVLI-APPPLLDWKVTOEKVPWGVILLLGGFALAKGSEASGL	429
QY	421	NIMISKAMQVLIGLPIVQSIITFVLNFFSAFNAVNVVANIPLILCEMSLALHPLI	480
Db	430	SVWKGQMEPLHAVPPAAITLILSLVAVFECTSNVATTLFLPIFASMSRSIGLNPLY	489
QY	481	LTPACLGISWYFLPVSTPPNAIVTOYAHIKTKYFACCGIVPTTIGISVALVNTWGL	540
Db	490	IMLPCTLSASFAMLPVATPPNAIVFTYGHKLVADMVKTVGMVNIIGVFCVFLAVNTWGR	549
QY	541	IIPFESKSPFDWA 553	
Db	550	AIF-DLDHFPDWA 561	

RESULT 8

ID	ADK51050	standard; protein; 568 AA.
XX	ADK51050;	
AC	ADK51050;	
DT	17-JUN-2004	(first entry)
XX	Human NOV18B	protein sequence SeqID70.
DE	cytostatic; NOVX-agonist; NOVX-antagonist; vaccine; gene therapy; cancer;	
XX	chromosome mapping; human; NOV18B.	
KW	Homo sapiens.	
XX	WO2003083046-A2.	
PN	09-OCT-2003.	
XX	01-APR-2003; 2003WO-US010142.	
PF	02-APR-2002; 2002US-00115479.	
XX	05-APR-2002; 2002US-0370349P.	
PR	08-APR-2002; 2002US-0370969P.	
PR	12-APR-2002; 2002US-0372019P.	
PR	22-APR-2002; 2002US-0374379P.	
PR	30-MAY-2002; 2002US-0384543P.	
PR	03-JUN-2002; 2002US-00160619.	
PR	15-AUG-2002; 2002US-0403748P.	

PR	04-NOV-2002; 2002US-00287226.	
PR	31-MAR-2003; 2003US-00403161.	
XX	(CURA-) CURAGEN CORP.	
PA	Anderson DW, Bento P, Boldog FL, Burgess CE, Casman SJ, Furtak K;	
PI	Gorman L, Gould-Rothberg BE, Gunther E, Heyes MP, Li L, Spytek KA,	
PI	Stone DJ, Zhong M, Malyankar UM, Edinger SR, Patturajan M;	
PI	Rothenberg ME, Smithson G;	
XX	WPI; 2003-812539/76.	
DR	N-PSDB; ADK51049.	
XX	New NOVX polypeptide, useful for preparing a composition for treating or	
PT	preventing e.g. cancer or for chromosome mapping.	
XX	Claim 1; SEQ ID NO 70; 433pp; English.	
XX	This invention relates to novel isolated polypeptides and the DNA	
CC	sequences which encode them. The invention may be useful for the	
CC	development of compounds with a cytostatic activity (as NOVX-agonists or	
CC	antagonists) or vaccines. In addition, the disclosed sequences may be	
CC	useful for gene therapy. The polypeptide is useful for preparing a	
CC	composition for treating or preventing a pathological state in a mammal,	
CC	for example cancer or for chromosome mapping. The present sequence is	
CC	that of a human NOVX protein of the invention.	
XX	Sequence 568 AA;	

Query Match 30.5%; Score 885; DB 7; Length 568;		Best Local Similarity 35.6%; Pred. No. 1.2e-82;	
Matches 197; Conservative 101; Mismatches 219; Indels 36; Gaps 9;			
QY	27	IIPLTLPILYIGFQDMAEFKLMIVTMALLWITETLPYVVTALFPLVFCPLGLVNA	86
Db	19	VTPLLLPLVI-----LMPAKVSCAVIILMAIYVCTEVIPLAVTSLMPVLLFPLQILDS	74
QY	87	SVVCKQYFDTIVVFLGGLIVALGIEYSNLHRIALRVIRVGGSPRRFLVGLMSVSTFM	146
Db	75	ROVCVQYKMDTNMLFLGGLIIVAVAVERNLHRIALRTLLMWGAKPARMLGFMGVFTALL	134
QY	147	GLWISNAGTAMCPIVKNVNELDN-----KIPFVYMTQEEBP-VEE	189
Db	135	SMWISNTATTAMVPIVEAILQOMEATSAATAGLELVDKGAKELPGSQVIFEGPTLQ	194
QY	190	GEPPHPSKITVAFYAGIAYASSIGGLTIGTGNLFRGIYTERPPTSTVETITFANFMF	249
Db	195	QEDQERKRLCKAWTLCICVYASIGGTATLTGTGNVLLGQNNELFPDSKDLVNFASWFA	254
QY	250	YSIPLMIVNVTLVIIAFL-ITHMGLFRPNSKTGKIIAEANTNRKLMEDVLRQRHIDLGP	308
Db	255	FAPFNNLV-----MLLFAWLQFVYMFSSFKKSGCGLESKKNKAALKVQEEYRKLGP	310
QY	309	MSCHIQMAIAFAFMIVLLITRKPGFVCGWSDLI-----NRKVGSASGLSFIVLILFALP	364
Db	311	LSFABINVLICFPLLVILWFSRDPGFMGWLTVAVVEGETKYVSDATVAIFVATLLFIVP	370
QY	365	TOYTFKVC-----GKGPFTAQDAIDALSWEYVURNPWGLLFLGGFALAVASREGL	420
Db	371	SQKPFNFRSQTEEGKSPVLI-APPPLLDWKVTOEKVPWGVILLLGGFALAKGSEASGL	429
QY	421	NIMISKAMQVLIGLPIVQSIITFVLNFFSAFNAVNVVANIPLILCEMSLALHPLI	480
Db	430	SVWKGQMEPLHAVPPAAITLILSLVAVFECTSNVATTLFLPIFASMSRSIGLNPLY	489
QY	481	LTPACLGISWYFLPVSTPPNAIVTOYAHIKTKYFACCGIVPTTIGISVALVNTWGL	540
Db	490	IMLPCTLSASFAMLPVATPPNAIVFTYGHKLVADMVKTVGMVNIIGVFCVFLAVNTWGR	549
QY	541	IIPFESKSPFDWA 553	
Db	550	AIF-DLDHFPDWA 561	

PA (MALY/) MALYANKAR U M.
PA (GERL/) GERLACH V.
PA (LIUY/) LIU Y.
PA (ANDE/) ANDERSON D W.
PA (SPAD/) SPADERNA S K.
PA (CATT/) CATTERTON E.
PA (LEIT/) LEITE M W.
PA (ZHON/) ZHONG H.
PA (ALSO/) ALSOBROOK J P.
PA (LEPL/) LEPLEY D M.
PA (RIEG/) RIEGER D K.
PA (BURG/) BURGESS C E.
XX
PI Padigaru M, Spytek KA, Shenoy SG, Taupier RJ, Pena CEA, Li L;
PI Zerrhuen BD, Gusev VI, Ji W, Gorman L, Miller CE, Kekuda R;
PI Patturajan M, Gangolli EA, Vernet CAM, Guo XS, Tchernev VI;
PI Fernandes ER, Casman SJ, Malyankar UM, Gerlach V, Liu Y;
PI Anderson DW, Spaderna SK, Catterton E, Leite MW, Zhong H;
PI Alsobrook JP, Lepley DM, Rieger DK, Burgess CE;
XX
DR WPI; 2004-225693/21.
DR N-PSDB; ADN61774.
XX
New NOVX polypeptides and nucleic acid molecules useful for diagnosing,
PT preventing or treating NOVX-associated disorders, e.g. cancer, diabetes,
PT infection or obesity, and in chromosome mapping, tissue typing or
PT pharmacogenomics.
XX
XX
XX Claim 1; SEQ ID NO 44; 786pp; English.
XX
XX The invention relates to an isolated polypeptide (designated NOVX, or
CC NOV1-NOV127) comprising a sequence selected from 178 fully defined amino
CC acid sequences (and their mature forms, variants and fragments). Also
CC included are an isolated nucleic acid molecule encoding NOVX, a vector
CC comprising the nucleic acid, a cell comprising the vector, methods for
CC determining the presence or amount of the polypeptide or the nucleic acid
CC molecule in a sample, methods for determining the presence of or
CC predisposition to a disease associated with altered levels of expression
CC of the above polypeptide or nucleic acid molecule in a first mammalian
CC subject, a method for identifying an agent that binds to the above
CC polypeptide, a method for identifying a potential therapeutic agent for
CC use in the treatment of a pathology that is related to aberrant
CC expression or physiological interactions of the polypeptide, a method of
CC screening for a modulator of activity or of latency or predisposition to
CC a pathology associated with the polypeptide and a method for modulating
CC the activity of the polypeptide cited above. The composition and methods
CC are useful for diagnosing, preventing or treating diseases such as
CC diabetes, obesity, infectious diseases, anorexia, cancer-associated
CC cachexia, cancer, neurodegenerative disorders like Alzheimer's disease or
CC Parkinson's disease, immune disorders, haematopoietic disorders,
CC dyslipidaemias, and other chronic diseases. These may also be used in
CC chromosome mapping, tissue typing, preventive medicine and
CC pharmacogenomics. The polypeptides are also useful as vaccines. The
CC present sequence represents a NOVX protein of the invention.
XX
XX Sequence 568 AA;
XX
XX Query Match 30.5%; Score 985; DB 8; Length 568;
XX Best Local Similarity 35.6%; Pred. No. 1.2e-82;
XX Matches 197; Conservative 101; Mismatches 219; Indels 36; Gaps 9;
XX
QY 27 IIPILITILYIGTDMABPKLWLVITMALLITETITPIYVITLPLVFCPLGLVNA 86
Db 19 VTPULLLPLVL-----LMPKRVSCAYVILMAIYCTEVIPLAVTSLMPVLLFPLFQILDS 74
QY 87 SIVCKQYFTDITVIVFLGLIYVALGIEYNSLHTRIALRVIRVGSRRLLFVGLMSVSTFM 146
Db 75 RQVCVQYMKDITNMLFLGLIYVAVVERWNLHRIALRTLLWVGAKPARLMGLFGMGVTALL 134
QY 147 GLWISNAGTAMCPIVKALVNEIDTN-----KIFPYMTQEEEP-VEE 189
Db 135 SMWISNTATTAMVPIVEAILQOMEATSAATEAGLELVDKGAKELPGSQVIFEGTILQ 194

QY 190 GEPPHPSKITVAFYAGIAYVASSIGGLTIGTGTNLVFRGIYTERPPTSTVEITFANFMF 249
Db 195 QEDQERKRLCKAWTLICICVAAISIGGTATLTGTGPNVLLGQNNELFPDPSKDLVNFASWFA 254
QY 250 YSIPLMWIVNVTLVIIAFL-ITHMGLFRNSKTKGIIAENANTRKLMEDVLRQRHIDLGP 308
Db 255 FAFPNMLV----MLLPFWLWLFQVYMFSSFKSGWCGLESKKEKAALKVLOEYKRLGP 310
QY 309 MSCHETQMAIAFAFMIVLLITRKPGFVPCGWSDLI-----NRKVVGSSASGLSFVLLIFALP 364
Db 311 LSFAEINVLICFLLVILWFSRDPGMPGWLIVAVVEGETKVSVDATVAIFVATLLFIYP 370
QY 365 TQYTFKYCC----GKGPFPTAQADAILSWEYVLRNIPWGLLFLGGGFALAVASRETGL 420
Db 371 SQKRFNFRSQTSEEGKSPVLI-APPPELLDWKVQEKVPMGIVLLILGGGFALAGSEASGL 429
QY 421 NMISKAMOVILGLNIVVQSTIFULANFFSFAFNANVVANIVLPILCMSLALEHPLI 480
Db 430 SYVMGKQMEPLHAVPPAAITLILSLVAVFTECTSNVATTTILFLPIFASMSRSIGLNPLY 489
QY 481 LTLPAICIGISMVYFLPVSTPPNAIVTQYAHIKTKYFACCGIVPTIIGISVALVNTNTWGL 540
Db 490 IMLPCTLSASFAMLPVATPPNAIVFTYCHLVADVMKVTGVMNIIIGVFCVLAVNTWGR 549
QY 541 IIFPESKSPFDWA 553
Db 550 AIF-DLDHFPDWA 561
XX
RESULT 11
ABR57023
ID ABR57023 standard; protein; 568 AA.
XX
AC ABR57023;
XX
XX 05-AUG-2003 (first entry)
DE Human TCH169 protein SEQ ID NO:1.
XX
XX Human; TCH169; dicarboxylate transport; hepatotropic; cytostatic;
XX nephrotropic; vasotropic; antidiabetic; liver disease; hepatitis;
XX hepatic sclerosis; alcohol-related liver disease; prostate disease;
XX prostatic disease; prostatic hypertrophy; spleen disease; spleen hyperactivity;
XX kidney disease; nephritis; kidney failure; nephritis; dropsy; diabetes;
XX diabetes-associated renal disease; metabolic disease; hyperlipaemia;
XX circulatory disease; arteriosclerosis; cancer.
XX
XX Homo sapiens.
XX WO2003025168-A1.
XX
XX 27-MAR-2003.
XX
XX 13-SEP-2002; 2002WO-JP009444.
XX
XX 17-SEP-2001; 2001JP-00281992.
XX 02-OCT-2001; 2001JP-00306873.
XX 16-APR-2002; 2002JP-00113279.
XX
XX (TAKE) TAKEDA CHEM IND LTD.
XX
XX Nakanishi A, Uno Y, Sagiya Y;
XX WPI; 2003-313352/30.
XX N-PSDB; ACC79539, ACC79558.
XX
XX Protein TCH169 with dicarboxylate transport activity for treatment and
XX diagnosis of diseases including liver disease, cancer and circulatory
XX disorders.
XX
XX Claim 1; Fig 1-2; 132pp; Japanese.
XX
XX The present invention describes protein TCH169 and its salts having
XX

dicarboxylate transport activity. TCH169 has hepatotropic, cytostatic, nephrotropic, vasotropic and antidiabetic activities. The TCH169 protein and polynucleotide can be used in the treatment, prevention and diagnosis of liver disease (such as hepatitis, hepatic sclerosis and alcohol-related liver disease); prostate disease (such as prostatitis and prostatic hypertrophy); spleen disease (such as splenic hyperactivity); kidney disease (such as nephritis, kidney failure, nephritis, dropsy and diabetes-associated renal disease); metabolic disease (such as diabetes) circulatory disease (such as hyperlipaemia and arteriosclerosis); and cancer (such as non-small cell lung cancer, liver cancer, renal cancer, ovarian cancer, prostate cancer, stomach cancer, pancreatic cancer, breast cancer, colon cancer, bladder cancer and womb cancer). The present sequence represents human TCH169, from the present invention

Query Match 30.2%; Score 875; DB 6; Length 568;
Best Local Similarity 35.7%, Pred. No. 1.3e-81;
Matches 198; Conservative 99; Mismatches 218; Indels 40; Gaps 10;

QY 27 IIPILITPILYGFQDMAEF-KCLMWLVTVMALLWITETLPDIYVTVAPLPVFCPLLLGNG 85
DB 19 VTPELLLPVLVI---LMPAKFRCAYVILLMAIYWCTEIVPLAVTSLMPVLLFPFLQILD 74
QY 86 ASIVCKQYFTDTIVVFGLGLVALGIEYSNHLTRIALRVIRVGSPRRRLFGVLMSTVF 145
DB 75 SRQCVCQYMKDTNNLFGLGLVAFAVERWNHLKRITALRTLTLWGAQPARLMLGPMGTAL 134
QY 146 MGLWISNSATGMCPVKALVNELDTN-----KIPIPVMTQEESP-VE 188
DB 135 LSMWISNTATTAMVPVIEAILQQMEATSAAATEAGLELVDKGKAKEUPGSQVIFEGPTLG 194
QY 189 EGEPHPHSKITVAFYAGYAIYASSIGGLTGTLTGTNLVFRGIYTERPPTSTVETTFANFM 248
DB 195 QOEQERKRCLKAMTLCICYAASIGGTATLTGTGNVLLQGMMLEFPDSKDLYNFASWP 254
QY 249 FYSIPLMVINVTUVIIAFLITHMGLFRPN-SKTGKIIAEANTNRKLMEDEVLRQRHIDL 307
DB 255 AFAPFNMLV---MLLPFWLWLQFYVMRFNFKSWGCGLESKKNEKAALKVQLQBEYRKLG 310
QY 308 PMSCHEIQMAIAFAMVILLITRKPGEVPGWSDLI----NRKVVSASGLSFIVLLIFAL 363
DB 311 PLSPAENVLICFFLLVILWFSDRGPMGPGLTVAWVEGETKYVSDATVAIFVATLLFI 370
QY 364 PTQYTFKYKC-----GKGPPTAQAIDAILSWEVLRNIWMGLLFLGGGFALAVASRET 418
DB 371 PSQXPKFNFSQTEEREKTPYP---PPLDWKVTQEKVPWGIVLLGGGFALAKGSEAS 427
QY 419 GLNIMI SKAQVLGLPNI VVQSITFVLANFFSAFNANVVVANIVLVPILCEMSIALSLHP 478
DB 428 GLSVMMGQMEPLHAVPPAAITLLSLVAVFTECTSNAVATTTLLFLPIFASMRSRIGLNP 487
QY 479 LILTLPACLGISMYFLPVSTPPNAIVTOAHIKTKYFACCGIVPTTIIGISVALVNTNW 538
DB 488 LYIMLPCTLGASFAFMLPVATPPNAIVFTYGHKLVDADMVKVTGIMNIIGVFCFVLAINTW 547
QY 539 GLIIFPSKSGFPDWA 553
DB 548 GRAIF-DLDHFDPWA 561

RESULT 12
AAE38764
ID AAE38764 standard; protein; 568 AA.
XX
AC AAE38764;
XX
XX 18-DEC-2003 (first entry)
XX
DE Human 69624 protein.
XX
KW Human; 69624; transporter protein; neurological disorder; therapy;
KW atherosclerosis; cardiac hypertrophy; ischaemia reperfusion injury;

CC useful in diagnosing, treating or preventing NOVA-associated disorders,
 CC e.g. cardiomyopathy, atherosclerosis, hypertension, cancer, obesity,
 CC diabetes, AIDS, multiple sclerosis, graft-versus-host disease,
 CC Alzheimer's disease, Parkinson's disease, asthma, or fertility disorders.
 CC The nucleic acids are further used as hybridization probes, in chromosome
 CC mapping, tissue typing, preventive medicine, and pharmacogenomics. The
 CC polypeptides are also useful as vaccines. This sequence represents an
 CC example of the polypeptide of the invention.
 XX
 SQ Sequence 568 AA;

Query Match 30.2%; Score 875; DB 8; Length 568;
 Best Local Similarity 35.7%; Pred. No. 1.3e-81;
 Matches 198; Conservative 99; Mismatches 218; Indels 40; Gaps 10;
 QY 27 IIPILITPLIYGFQDTMAEF-KCLWLIVTMALLWITETPIYVVTALFLVFPCLGLV 85
 DB 19 VTPLLLPLVI-----LMPAKFVRCAYIILMAIYWCTEVPLAVTSLMPVLLPQLD 74
 QY 86 ASIVCKQYFDTIIVFLGGLIIVAGIEYSNLHTRIALRVIRIVGSGPRRLFVGLMSVSTF 145
 DB 75 SRQVCVQYKMDTNLFLGGLIIVAVVERNLHTRIALRVIRIVGSGPRRLFVGLMSVSTF 134
 QY 146 MGLWISNSAGTAMCPIVKALVNELDN-----KIPFVYMTQEEEP-VE 188
 DB 135 LSMWISNTATTAMVPIVEAILQOMEATSAATEAGLELVDKGAKEPGLSQVIFEGPTLG 194
 QY 189 EGGPPHPSKITVAFYAGIAYASSIGGLTGTLTGNTLVFRGIYTERPPTSTVITANFM 248
 DB 195 QBEDQERKRLCKAMTLCICVAAISGGTATLTGTGPNVLLGQMNELFPDSDKLVNFASF 254
 QY 249 FYSIPLMVIVNVTIIVFLGGLIIVAGIEYSNLHTRIALRVIRIVGSGPRRLFVGLMSVSTF 307
 DB 255 AFAPNNLV-----MLLPFWLWLFQVYVNRFNFKKSWGGCLESKKNEKALKVLQEEYRKL 310
 QY 308 PMSCHEIQMAIAFAFMIVLLITRKPGFVPGWSDLI-----NRKVGSAGSLFIVLLIFAL 363
 DB 311 PLPSFAEINVLICFLLVILWFSRDPGMPGWLTVAWVEGETKVVSDATVAIFVATLLFIV 370
 QY 364 PTQYTFKYYCC-----GKGFPTAQADAILSWYEVLRNIPWGLLFLGCGFALAKSEAS 418
 DB 371 PSQKPFNFRSQTEERKTFYP-----PPLLDWKVTQEKVPGWVLLGGGFALAKSEAS 427
 QY 419 GLNIMISKAMQVLIGLNPVQSIITFVLNFFSAFNANVVVANIPLCEMSLALHLP 478
 DB 428 GLSVWKGQNEPLHVPAPPAITLILSLVAVFECTSNVATTLFLFIPASMSRSIGLNP 487
 QY 479 LILTPACLGISMVYFLPVSTPPNAIVTQYAHIKTKYFACCGIVPTIIGISVALVNTW 538
 DB 488 LYIMLPCTLSASFAPMLPVATPPNAIVFTYGHKLKVDAMVKTGVTIMNIIGVFCVFLAVNTW 547
 QY 539 GLIIPESKSPDDWA 553
 DB 548 GRAIF-DLDHFPDWA 561

RESULT 15

ID ADP64793
 AC ADP64793 standard; protein; 568 AA.
 AC ADP64793;
 XX
 XX
 DT 26-AUG-2004 (first entry)
 XX Human Na⁺-coupled citrate transporter protein.
 DE sodium-coupled citrate transporter; transmembrane citrate transporter;
 XX lifespan; weight reduction; weight gain prevention; blood cholesterol;
 KW triglyceride; low density lipopolysaccharide; glucose; obesity;
 KW hyperlipidemia; hypercholesterolemia; INDY protein.
 OS Homo sapiens.
 XX

PN WO2004048925-A2.
 XX 10-JUN-2004.
 XX 20-NOV-2003; 2003WO-US037054.
 XX 22-NOV-2002; 2002US-0428469P.
 PR 01-APR-2003; 2003US-0459441P.
 XX (MEDI-) MEDICAL COLLEGE GEORGIA RES INST.
 PA (GANA/) GANAPATHY V.
 PA (INOU/) INOUE K.
 PA (FEIY/) FEI Y.
 XX Ganapathy V, Inoue K, Fei Y;
 XX WPI; 2004-460797/43.
 DR N-PSDB; ADP64792.
 XX
 PT New isolated polynucleotide encoding a Na⁺-coupled citrate transporter
 (NaCT) polypeptide, useful as a drug target for the treatment of obesity,
 PT hyperlipidemia, and hypercholesterolemia.
 XX Claim 44; SEQ ID NO 6; 186pp; English.
 XX The invention relates to novel Na⁺-coupled citrate transporter proteins
 and their encoding genes. Inhibitors of transmembrane citrate
 CC transporters are useful for extending the lifespan, reducing weight,
 CC preventing weight gain or lowering blood cholesterol, triglyceride, LDL
 CC or glucose levels in a subject. The NaCT polypeptide is useful as a drug
 CC target for the treatment of obesity, hyperlipidemia, and
 CC hypercholesterolemia. This sequence corresponds to the human Na⁺-coupled
 CC citrate transporter protein.
 XX
 SQ Sequence 568 AA;
 Query Match 30.2%; Score 875; DB 8; Length 568;
 Best Local Similarity 35.7%; Pred. No. 1.3e-81;
 Matches 198; Conservative 99; Mismatches 218; Indels 40; Gaps 10;
 QY 27 IIPILITPLIYGFQDTMAEF-KCLWLIVTMALLWITETPIYVVTALFLVFPCLGLV 85
 DB 19 VTPLLLPLVI-----LMPAKFVRCAYIILMAIYWCTEVPLAVTSLMPVLLPQLD 74
 QY 86 ASIVCKQYFDTIIVFLGGLIIVAGIEYSNLHTRIALRVIRIVGSGPRRLFVGLMSVSTF 145
 DB 75 SRQVCVQYKMDTNLFLGGLIIVAVVERNLHTRIALRVIRIVGSGPRRLFVGLMSVSTF 134
 QY 146 MGLWISNSAGTAMCPIVKALVNELDN-----KIPFVYMTQEEEP-VE 188
 DB 135 LSMWISNTATTAMVPIVEAILQOMEATSAATEAGLELVDKGAKEPGLSQVIFEGPTLG 194
 QY 189 EGGPPHPSKITVAFYAGIAYASSIGGLTGTLTGNTLVFRGIYTERPPTSTVITANFM 248
 DB 195 QBEDQERKRLCKAMTLCICVAAISGGTATLTGTGPNVLLGQMNELFPDSDKLVNFASF 254
 QY 249 FYSIPLMVIVNVTIIVFLGGLIIVAGIEYSNLHTRIALRVIRIVGSGPRRLFVGLMSVSTF 307
 DB 255 AFAPNNLV-----MLLPFWLWLFQVYVNRFNFKKSWGGCLESKKNEKALKVLQEEYRKL 310
 QY 308 PMSCHEIQMAIAFAFMIVLLITRKPGFVPGWSDLI-----NRKVGSAGSLFIVLLIFAL 363
 DB 311 PLPSFAEINVLICFLLVILWFSRDPGMPGWLTVAWVEGETKVVSDATVAIFVATLLFIV 370
 QY 364 PTQYTFKYYCC-----GKGFPTAQADAILSWYEVLRNIPWGLLFLGCGFALAKSEAS 418
 DB 371 PSQKPFNFRSQTEERKTFYP-----PPLLDWKVTQEKVPGWVLLGGGFALAKSEAS 427
 QY 419 GLNIMISKAMQVLIGLNPVQSIITFVLNFFSAFNANVVVANIPLCEMSLALHLP 478
 DB 428 GLSVWKGQNEPLHVPAPPAITLILSLVAVFECTSNVATTLFLFIPASMSRSIGLNP 487
 QY 479 LILTPACLGISMVYFLPVSTPPNAIVTQYAHIKTKYFACCGIVPTIIGISVALVNTW 538

Db 488 LYIMLPCTLSASFAPMLPVATPPNAIVFTYGHKLVADMYKTGVIMNIIIGVFCVFLAVNTW 547
QY 539 GLIIFPEKSFPDWA 553
Db 548 GRAIF-DLDHFPDWA 561

Search completed: June 30, 2005, 08:52:04
Job time : 82.2935 secs

THIS PAGE IS BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2005, 08:36:57 ; Search time 21.608 Seconds
(without alignments)
1941.542 Million cell updates/sec

Title: US-10-017-479A-3

Perfect score: 2897

Sequence: 1 MAEPGEQRKFLVLRCCIFHW.....FPESKSPDWAKEIKNQTKI 562

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/prodata/1/aaa/5A COMB.pep:*
- 2: /cgn2_6/prodata/1/aaa/5B COMB.pep:*
- 3: /cgn2_6/prodata/1/aaa/6A COMB.pep:*
- 4: /cgn2_6/prodata/1/aaa/6B COMB.pep:*
- 5: /cgn2_6/prodata/1/aaa/PTUS COMB.pep:*
- 6: /cgn2_6/prodata/1/aaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	920	31.8	180	4	US-09-270-767-42669
2	844	29.1	599	4	US-09-549-016-9866
3	793	27.4	368	4	US-09-270-767-42029
4	743	25.6	626	4	US-09-556-916-20
5	742.5	25.6	627	4	US-09-556-916-8
6	741	25.6	626	4	US-09-556-916-14
7	740.5	25.6	627	4	US-09-556-916-2
8	738.5	25.5	627	4	US-09-549-016-6840
9	737	25.4	626	4	US-09-556-916-22
10	736.5	25.4	627	4	US-09-556-916-10
11	735	25.4	626	4	US-09-556-916-16
12	734.5	25.4	627	4	US-09-556-916-4
13	696	24.0	601	4	US-09-949-016-9977
14	661	22.8	132	4	US-09-270-767-57987
15	624	21.5	580	4	US-09-556-916-24
16	623.5	21.5	581	4	US-09-556-916-12
17	622	21.5	580	4	US-09-556-916-18
18	621.5	21.5	581	4	US-09-556-916-6
19	611	21.1	230	4	US-09-270-767-43713
20	582.5	20.1	561	4	US-09-949-016-8161
21	544	18.8	527	4	US-09-602-787A-516
22	527	18.2	194	4	US-09-270-767-59102
23	453.5	15.7	470	4	US-09-543-681A-5952
24	427.5	14.8	169	4	US-09-270-767-57286
25	351	12.1	233	4	US-09-489-847-176
26	327.5	11.3	335	4	US-09-602-787A-518
27	283.5	9.8	524	3	US-09-134-001C-5457

Sequence 20749, A
Sequence 4343, Ap
Sequence 12605, A
Sequence 10445, A
Sequence 2981, Ap
Sequence 287, App
Sequence 6984, Ap
Sequence 22637, A
Sequence 532, App
Sequence 13843, A
Sequence 4729, Ap
Sequence 5201, Ap
Sequence 11870, A
Sequence 7512, Ap
Sequence 13547, A
Sequence 5886, Ap
Sequence 58, Appl
Sequence 5065, Ap
Sequence 7033, Ap
Sequence 6790, Ap
Sequence 2, Appl
Sequence 4, Appl
Sequence 4637, Ap
Sequence 2120, Ap
Sequence 25488, A
Sequence 9836, Ap
Sequence 6105, Ap
Sequence 4523, Ap
Sequence 7784, Ap
Sequence 6084, Ap
Sequence 4250, Ap
Sequence 18036, A
Sequence 10612, A
Sequence 5049, Ap
Sequence 7401, Ap
Sequence 42613, A
Sequence 6031, Ap
Sequence 209, App
Sequence 5213, Ap
Sequence 5864, Ap
Sequence 8240, Ap
Sequence 10570, A
Sequence 5510, Ap
Sequence 10360, A
Sequence 13589, A
Sequence 32040, A
Sequence 12980, A
Sequence 7515, Ap
Sequence 20740, A
Sequence 4309, Ap
Sequence 5185, Ap
Sequence 4822, Ap
Sequence 444, App
Sequence 12, Appl
Sequence 284, App
Sequence 3526, Ap
Sequence 11893, A
Sequence 236, App
Sequence 7690, Ap
Sequence 3001, Ap
Sequence 3355, Ap
Sequence 3355, Ap
Sequence 4808, Ap
Sequence 568, Ap
Sequence 60, Appl
Sequence 4582, Ap
Sequence 4997, Ap
Sequence 7336, Ap
Sequence 6008, Ap
Sequence 5089, Ap
Sequence 20407, A
Sequence 7497, Ap

ALIGNMENTS

RESULT 1

US-09-270-767-42669
; Sequence 42669, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 42669
; LENGTH: 180
; TYPE: PRT
; ORGANISM: *Drosophila melanogaster*
US-09-270-767-42669

Query Match 31.8%; Score 920; DB 4; Length 180;
Best Local Similarity 100.0%; Pred. No. 1.3e-84;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 383 AIDAILSWYVLRNIPWGLLFLGGGFALAVASRETGLNIMISKAMQVILGLNIVVQSI 442
DB 1 AIDAILSWYVLRNIPWGLLFLGGGFALAVASRETGLNIMISKAMQVILGLNIVVQSI 60

QY 443 TFVLNFFSAPNANVVANIVLPILCENSLALELHPLILTPACLGISMVYELPVPSTPPN 502
DB 61 TFVLNFFSAPNANVVANIVLPILCENSLALELHPLILTPACLGISMVYELPVPSTPPN 120

QY 503 AIVTQYAHIKTKYPACCGIVPTIIGISVALVNTNTWGLIIFPESKSPDPWAKEIKNOTKI 562
DB 121 AIVTQYAHIKTKYPACCGIVPTIIGISVALVNTNTWGLIIFPESKSPDPWAKEIKNOTKI 180

RESULT 2

US-09-949-016-9866
; Sequence 9866, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9866
; LENGTH: 599
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-9866

Query Match 29.1%; Score 844; DB 4; Length 599;
Best Local Similarity 33.9%; Pred. No. 3.5e-76;
Matches 200; Conservative 111; Mismatches 199; Indels 80; Gaps 14;

QY 20 WRGKASII---IPLITPILIIYGFQTDWAEFKCLWLIVTALLWITETLPIYVTALEPLV 76
DB 16 WAYRSYLIVFVFPILLPLPIL---VPSKEAYCAYAILMALFWCTEALPLAVTALPFLI 72

QY 77 FCPLLGLVNASIVCKQYFDTTIVVFLGGLIVLALGIEYSNLHTRIALRVIRIVGSGRRLLF 136
DB 73 LFPWGVIVDASEVAVEYLDKSNLLFFGGLLVAIAVEHNLHKEIALRVLLIVGVRAPLI 132

QY 137 VGLMSVSTFMGLWINSAGTAMMCPIVKALVNELDNKNIPFVYMTQEEBPVEBG----- 190
DB 133 LGFMLVTAFLSMWISNTATSAMVPIAHAVLDQL-----HSSQASSNVEGSGNNPTF 184

QY 191 --EPHPHSK-----ITVAFYAG-----IYASSIGIGLTL 218
DB 185 ELQEPSQKEVTKLDNGQALPVTSSASSEGRAHLSQKHLHLTQCMSCVCYSASIGGLATL 244

QY 219 IGTGTLNLFVRGIYTERFPTSTVEITPANEMFYISPLMWIVNVTLVIAFL---ITHMGL- 274
DB 245 TGTAPNLVLQGINSLFPQNGVNVNPFASFPTWVI-----LLLAWLWLQILFLGN 300

QY 275 FRNSKTGKIIABANTNRKLMEDVLRQRHIDIGPMSCHEIQMAIAFAFMVLVLTITKPGF 334
DB 301 FRKNFGIGEKMQE---QOQAAVCVIQTEHRLGPMPTFAEKASISILEFVILVLLWFTREPGF 357

QY 335 VPCWSDLI-----NRKVVGSAGLSFVILLIPALPTQYTFYFKYCCG-----KGPFTAAAI 384
DB 358 FLGWNGLAPFNKAGESWSDGTVAIFIGIIMFIIPSKFP-----GLTQDPENPGKLKAP 411

QY 385 DAILSWYVLRNIPWGLLFLGGGFALAVASRETGLNIMISKAMQVILGLNIVVQSI 444
DB 412 LGLLDWKTQKMPNIVLLGGGYALAKGSGRSLSEWLGKLTPLQSVAPAPAIILS 471

QY 445 VLANFFSAPNANVVANIVLPILCENSLALELHPLILTPACLGISMVYELPVPSTPPN 504
DB 472 LLVATFECTSNVATTTIFLPLASMAQAICLHPLYVMLPCTLATSLAFMLPVATPPN 531

QY 505 VITQYAHIKTKYPACCGIVPTIIGISVALVNTNTWGLIIFPESKSPDPWAK 554
DB 532 VFSFGDLKVDMARAGFLNIIGVLIILAINSWGIPLP-SLHSFFSWAQ 580

RESULT 3

US-09-270-767-42029
; Sequence 42029, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 42029
; LENGTH: 368
; TYPE: PRT
; ORGANISM: *Drosophila melanogaster*
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-42029

Query Match 27.4%; Score 793; DB 4; Length 368;
Best Local Similarity 43.6%; Pred. No. 2.4e-71;
Matches 154; Conservative 71; Mismatches 112; Indels 16; Gaps 6;

QY 3 EPGEQRKFVLGRCCIF---HWRGKASIIIPLTITPILIIYGFQTDMAEFKCLWLIVTALL 59
DB 17 EIGEQQPPV-KCSNFFANHWKGLVFLVLLCLPVMLL--NEGAEFRCLWLLVMAIF 72

QY 60 WITETLPIYVTALEPLVFCPLLGLVNASIVCKQYFDTTIVVFLGGLIVLALGIEYSNLHTR 119
DB 73 WVTETALPLYVTSIMPIVAFPIIMGIMSSDQTCRLYFKDTLVFMFGGINVALAVEYCNLHKR 132

QY 120 IALRVIRIVGSGRRLLFVGLMSVSTFMGLWINSAGTAMMCPIVKALVNELDNKNIPFV 179
DB 133 IALRVIRIVGSGRRLLFVGLMSVSTFMGLWINSAGTAMMCPIVKALVNELDNKNIPFV 192

Qy 479 LILTPACLGISMVYFLPVPSTPPNAIVTQYAHIKTKYFACCGVPTTIIGISVALVNTNW 538
Db 545 LYTLIPVTMCISFAVMLPVGNPPNAIVFSYGHQCIKDMVKAGLVNVLGLVVMVAINTW 604
Qy 539 GLIIFPESKSPDWAKEIKNOT 560
Db 605 GVSLF-HLDTYPAWAR-VSNIT 624

RESULT 6
US-09-556-916-14
; Sequence 14, Application US/09556916
; Patent No. 6548271
; GENERAL INFORMATION:
; APPLICANT: Turner, Alex
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins
; FILE REFERENCE: 8535-041-999
; CURRENT APPLICATION NUMBER: US/09/556,916
; CURRENT FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-556-916-14

Query Match 25.6%; Score 741; DB 4; Length 626;
Best Local Similarity 28.2%; Pred. No. 9.1e-66;
Matches 175; Conservative 123; Mismatches 223; Indels 100; Gaps 14;

Qy 26 IILPLITLPIILYGFQTDMAEFKCLMLIVTMALLMITETLPIYVTALFPLVFCPLGLV 85
Db 17 VCVPLLLPLVL---HPSEASCAVLIIVTAVVWSEAVPLGAAALVPAFLYFFFGVLR 73
Qy 86 ASIVCKQYFTDTIVFVGLGILVALGIEYSNLHTRIALRVIRIVGGSPRRLFVGLMSVSTF 145
Db 74 SNEVAEYFKNNTLLLVGVICVAAVEKNLHRIALRMVLMAGAKPGMLLLCFMCTTL 133
Qy 146 MGLWISNSAGTAMCPVICALVNEL---DTNKIFPVYMTQEEPV-----187
Db 134 LSMWLSNTSTTAMVPIVEAVLQELVSAEDEQLVAGNSNTEEAEPISLDVNSQPSLELI 193
Qy 188 -----EGEPHPHSK-----197
Db 194 FVNEESNADLTLMHNENLVGPSITNPIKTANQHCKQHPSQEKPVLTSPRQKLN 253
Qy 198 -----ITVAFYAGIAYASSIGGLTGLTGNLVFRGIYTERPPTSTVEITFANF 247
Db 254 RKYRSHHDOMICKLSLSISYSATIGLTTIGTSTSLIFLEHFNNOYPAAEV-VNFGTW 312
Qy 248 MFYSIPLMVNVTLVIAFLITHMGLFRPNSK-TGKIIAEANTNR-KLMEDVLRQHD 305
Db 313 FLFSFPISLI---MLVVSFWFMHWLFLGCFNFKETCSLKKKTKRQLSEKRIQEEYK 368
Qy 306 LGPMSCHIEIQMAIAFAFMIVLLITRKPGFVPGWSDLINRNVKVSASGLS-FIVLLIFALP 364
Db 369 LGDISYPEMVTGPFILMTVLMFTREPGFVPGWDSFPEKKGYRTDATVSVFLGFLFLIP 428
Qy 365 TQYTFPKYCCGK---GPFTAQAI---DAILSWEYVLRNIPWGLLFLGGGFALAVASRETG 419
Db 429 AK-----KPCFGKNDGENQEHSLGTETPIITWKDFQKTPWEIVILVGGGVALASGSKSSG 484
Qy 420 LNMISKAMQVLIGLPNIIVQSITFVLNFFSAFNANVVANVIVLPILCEMSLALHPL 479
Db 485 LSTWIGNQMLSSLSLPPWATLLACILVSIVTEFVSNPATITIFLPILCSLSETLHNP 544
Qy 480 ILTLPACLGISMVYFLPVPSTPPNAIVTQYAHIKTKYFACCGVPTTIIGISVALVNTNW 539

Db 545 YTLIPVTMCISFAVMLPVGNPPNAIVFSYGHQCIKDMVKAGLVNVLGLVVMVAINTW 604
Qy 540 LIIFPESKSPDWAKEIKNOT 560
Db 605 VSLF-HLDTYPAWAR-VSNIT 623

RESULT 7
US-09-556-916-2
; Sequence 2, Application US/09556916
; Patent No. 6548271
; GENERAL INFORMATION:
; APPLICANT: Turner, Alex
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins
; FILE REFERENCE: 8535-041-999
; CURRENT APPLICATION NUMBER: US/09/556,916
; CURRENT FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 627
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-556-916-2

Query Match 25.6%; Score 740.5; DB 4; Length 627;
Best Local Similarity 28.1%; Pred. No. 1e-65;
Matches 175; Conservative 123; Mismatches 223; Indels 101; Gaps 14;

Qy 26 IILPLITLPIILYGFQTDMAEFKCLMLIVTMALLMITETLPIYVTALFPLVFCPLGLV 85
Db 17 VCVPLLLPLVL---HPSEASCAVLIIVTAVVWSEAVPLGAAALVPAFLYFFFGVLR 73
Qy 86 ASIVCKQYFTDTIVFVGLGILVALGIEYSNLHTRIALRVIRIVGGSPRRLFVGLMSVSTF 145
Db 74 SNEVAEYFKNNTLLLVGVICVAAVEKNLHRIALRMVLMAGAKPGMLLLCFMCTTL 133
Qy 146 MGLWISNSAGTAMCPVICALVNEL---DTNKIFPVYMTQEEPV-----187
Db 134 LSMWLSNTSTTAMVPIVEAVLQELVSAEDEQLVAGNSNTEEAEPISLDVNSQPSLELI 193
Qy 188 -----EGEPHPHSK-----197
Db 194 FVNEESNADLTLMHNENLVGPSITNPIKTANQHCKQHPSQEKPVLTSPRQKLN 253
Qy 198 -----ITVAFYAGIAYASSIGGLTGLTGNLVFRGIYTERPPTSTVEITFAN 246
Db 254 RKYRSHHDOMICKLSLSISYSATIGLTTIGTSTSLIFLEHFNNOYPAAEV-VNFGT 312
Qy 247 MFYSIPLMVNVTLVIAFLITHMGLFRPNSK-TGKIIAEANTNR-KLMEDVLRQHI 304
Db 313 FLFSFPISLI---MLVVSFWFMHWLFLGCFNFKETCSLKKKTKRQLSEKRIQEEYE 368
Qy 305 LGPMSCHIEIQMAIAFAFMIVLLITRKPGFVPGWSDLINRNVKVSASGLS-FIVLLIFAL 363
Db 369 LGDISYPEMVTGPFILMTVLMFTREPGFVPGWDSFPEKKGYRTDATVSVFLGFLFLI 428
Qy 364 TQYTFPKYCCGK---GPFTAQAI---DAILSWEYVLRNIPWGLLFLGGGFALAVASRET 418
Db 429 PAK-----KPCFGKNDGENQEHSLGTETPIITWKDFQKTPWEIVILVGGGVALASGSKSS 484
Qy 419 GLNMTISKAMQVLIGLPNIIVQSITFVLNFFSAFNANVVANVIVLPILCEMSLALHLP 478
Db 485 GLSTWIGNQMLSSLSLPPWATLLACILVSIVTEFVSNPATITIFLPILCSLSETLHNP 544
Qy 479 LILTPACLGISMVYFLPVPSTPPNAIVTQYAHIKTKYFACCGVPTTIIGISVALVNTNW 538
Db 545 LYTLIPVTMCISFAVMLPVGNPPNAIVFSYGHQCIKDMVKAGLVNVLGLVVMVAINTW 604


```

Db      605 VSLF-HLDTYPAPAWAR-VSNIT 623

RESULT 10
US-09-556-916-10
; Sequence 10, Application US/09556916
; Patent No. 6548271
; GENERAL INFORMATION:
; APPLICANT: Turner, Alex
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins
; FILE REFERENCE: 8535-041-999
; CURRENT APPLICATION NUMBER: US/09/556,916
; CURRENT FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 627
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-556-916-10

Query Match      25.4%; Score 736.5; DB 4; Length 627;
Best Local Similarity 28.0%; Pred. No. 2.6e-65;
Matches 174; Conservative 124; Mismatches 223; Indels 101; Gaps 14;

Qy      26 IIPILITLPIIYGFQTDMAEFKCLMLIVTMALLMITETLPIYVYVTFPLVFCPLLLGLVN 85
Db      17 VCVPLLLLPVL---HPSSEASCAVLIIVTAVYVWSEAVPLGAAALVPAPFLYPPFGVLR 73
Qy      86 ASIVCKQYFTDITVIFGLGIVAGIYSNLHRIALRVIRIVGSPRRFLVGLMSVSTF 145
Db      74 SNEVAEYFKNTLLLVGVICVAAAEKVNHLKRIALRMVLMAGAKPGMLLLCFMCCCTL 133
Qy      146 MGLWISNSAGTAMCPIVKALVNEL-----DTNKIPFVYMTQEEPV----- 187
Db      134 LSMWLSNTSTTAMVMPIVEAVLQELVSAEDEQLVAGNSNTEAEPSLVDKNSQPSLELI 193
Qy      188 -----EAGEPPHPSK----- 197
Db      194 FVNEDRSNADLTLMHNENLNGVPSITNPIKTANQHQKQHPSEKQPVLTPSPRKQLN 253
Qy      198 -----ITVAFVAGIAYASSIGGLTGLTGTGNLVFRGIYTERFPTSTVEITFAN 246
Db      254 NRKYRSHHDQMI CKLSLSISYSATIGGLTIIIGTSTLSLIFLHFNNQYPAAEV-VNFGT 312
Qy      247 FMFYSIPLMIVNVTVIIAFLITHMGLFRPNSK-TGKIIAEANTNR-KLMEDVLRQRHI 304
Db      313 WFLFSPFISLI---MLVVSFWFMHMLFLGCNFKETCSLSKKKTKREQLSEKRIQEEYE 368
Qy      305 DLGPMSCHEIOMAIAPAFMVLITRKPGVPWGSDDLNRKVVGSASGLS-FIVLLIFAL 363
Db      369 KLGDISPEMVGTGFFILMTVLTREPFGVPWGSDFEKKGYRTDATVSFVLFGLLELI 428
Qy      364 PTQYTFPKYCCGK---GPFTAQAI---DAILSWEVYLRNIPWGLLFLGGGFALAVASRET 418
Db      429 PAK-----KPCFGKKNQDGENQEHSLGTSIITWKDFQKTMPEWIVLVGGGYALASGSKS 484
Qy      419 GLNIMISKAMQVLIGLPNIYVQSITFVLNFFSAFNANVVVANIVLPILCMSLALHLHP 478
Db      485 GLSTWIGNQMLSSLPWPWATLLACLIVSIVTFVSNPATITIFILPILCSLSETOHNP 544
Qy      479 LIILTPACLGISMVYFLPVSTPPNAIVTQYAHIKTKYFACCGIVPTIIGISVALVNTNW 538
Db      545 LYTLIPVTMCISFAVMLPVGNPNNAIVFSYGHQCIKDMVYAGLVGNVIGLIVVMVAINTW 604
Qy      539 GLIIFPESKFPDPAWEIKNQ 560
Db      605 GVSLF-HLDTYPAPAWAR-VSNIT 624

RESULT 11
US-09-556-916-16
; Sequence 16, Application US/09556916
; Patent No. 6548271
; GENERAL INFORMATION:
; APPLICANT: Turner, Alex
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins
; FILE REFERENCE: 8535-041-999
; CURRENT APPLICATION NUMBER: US/09/556,916
; CURRENT FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-556-916-16

Query Match      25.4%; Score 735; DB 4; Length 626;
Best Local Similarity 28.0%; Pred. No. 3.6e-65;
Matches 174; Conservative 123; Mismatches 224; Indels 100; Gaps 14;

Qy      26 IIPILITLPIIYGFQTDMAEFKCLMLIVTMALLMITETLPIYVYVTFPLVFCPLLLGLVN 85
Db      17 VCVPLLLLPVL---HPSSEASCAVLIIVTAVYVWSEAVPLGAAALVPAPFLYPPFGVLR 73
Qy      86 ASIVCKQYFTDITVIFGLGIVAGIYSNLHRIALRVIRIVGSPRRFLVGLMSVSTF 145
Db      74 SNEVAEYFKNTLLLVGVICVAAAEKVNHLKRIALRMVLMAGAKPGMLLLCFMCCCTL 133
Qy      146 MGLWISNSAGTAMCPIVKALVNEL-----DTNKIPFVYMTQEEPV----- 187
Db      134 LSMWLSNTSTTAMVMPIVEAVLQELVSAEDEQLVAGNSNTEAEPSLVDKNSQPSLELI 193
Qy      188 -----EAGEPPHPSK----- 197
Db      194 FVNEDRSNADLTLMHNENLNGVPSITNPIKTANQHQKQHPSEKQPVLTPSPRKQLN 253
Qy      198 -----ITVAFVAGIAYASSIGGLTGLTGTGNLVFRGIYTERFPTSTVEITFAN 247
Db      254 NRKYRSHHDQMI CKLSLSISYSATIGGLTIIIGTSTLSLIFLHFNNQYPAAEV-VNFGTW 312
Qy      248 FMFYSIPLMIVNVTVIIAFLITHMGLFRPNSK-TGKIIAEANTNR-KLMEDVLRQRHI 305
Db      313 WFLFSPFISLI---MLVVSFWFMHMLFLGCNFKETCSLSKKKTKREQLSEKRIQEEYE 368
Qy      306 LGPMSCHETOMAIAPAFMVLITRKPGVPWGSDDLNRKVVGSASGLS-FIVLLIPALP 364
Db      369 KLGDISPEMVGTGFFILMTVLTREPFGVPWGSDFEKKGYRTDATVSFVLFGLLELI 428
Qy      365 TQYTFPKYCCGK---GPFTAQAI---DAILSWEVYLRNIPWGLLFLGGGFALAVASRET 419
Db      429 AK-----KPCFGKKNQDGENQEHSLGTEPIITWKDFQKTMPEWIVLVGGGYALASGSKS 484
Qy      420 LNMISKAMQVLIGLPNIYVQSITFVLNFFSAFNANVVVANIVLPILCMSLALHLHP 479
Db      485 LSTWIGNQMLSSLPWPWATLLACLIVSIVTFVSNPATITIFILPILCSLSETOHNP 544
Qy      480 IILTLPACLGISMVYFLPVSTPPNAIVTQYAHIKTKYFACCGIVPTIIGISVALVNTNW 539
Db      545 YTLIPVTMCISFAVMLPVGNPNNAIVFSYGHQCIKDMVYAGLVGNVIGLIVVMVAINTW 604
Qy      540 LIIFPESKFPDPAWEIKNQ 560
Db      605 VSLF-HLDTYPAPAWAR-VSNIT 623
```

RESULT 12

US-09-556-916-4
; Sequence 4, Application US/09556916
; Patent No. 6548271
; GENERAL INFORMATION:
; APPLICANT: Turner, Alex
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins
; CURRENT APPLICATION NUMBER: US/09/556,916
; FILING DATE: 8535-041-999
; CURRENT FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 627
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-556-916-4

Query Match 25.4%; Score 734.5; DB 4; Length 627;
Best Local Similarity 28.0%; Pred. No. 4.1e-65;
Matches 174; Conservative 123; Mismatches 224; Indels 101; Gaps 14;
QY 26 IIPITLPIIYGFQTDMAEFCGLWLVITMALLMITETLPYVTVLPLVFCPLGLLVN 85
DB 17 VCVPLLLPLVL---HPSSCAVVLIVTAVVWSEAVPLGAALVPAFLYFFGVLR 73
QY 86 ASIVCKQFTDIIVFLGLIIVGALGIEYSLHTRIALRVIRVGGSPRRFLVGLMSVTF 145
DB 74 SNEVAEYFKNITLLVGVICVAAVEKWNHKLRLMVLMAKAPGMLLFCMCCTTL 133
QY 146 MGLWISNSAGTAMCPYKVALNEL---DTNKIFPVYMTQEEV----- 187
DB 134 LSWLSNISTTANVMPVEAVLQELVSADEOLVAGNSNTEBAEPLSLDVKNQPSLEI 193
QY 188 -----EEGEPHPSPK----- 197
DB 194 FVNEDRSNADLTLMHNENLNGVPSITNPIKTANQHQKQHPQSEKQVLTSPSPKOKL 253
QY 198 -----ITVAFYAGIYVASSIGGLTGLTGNLVRFGIYTERPPTSTVEITPAN 246
DB 254 NRKYRSHHDQMIKCLSLSSISYSATIGLTIIGTSTLSLPLEHFNQYPAEV-VNEG 312
QY 247 PMFYSIPLMVIIVNVLIIAFELIHMGLFRPNSK-TGKIIAEANTNR-KLMEDVLQRHI 304
DB 313 WFLSFPISLI---MLVVSFWHMLFLGNCNFKETCSLSKSKKTKRQSEKRIQEEYE 368
QY 305 DLGPMSCHEIQMAIAPFMIIVLLITRKPGFVPGWSDNLNRKVVGSASGLS-FIVLLIFAL 363
DB 369 KLGDISYEMVTGFFILMTVLMFTREGFVPGWSDSFEKKGRTDATVSVFLGLFLI 428
QY 364 PQYTFYFYCCGK---GPTTAQAI---DAILGWEYVLRNIPMGLLFLGCGFALAVASRET 418
DB 429 PAK---KPCFGKNDGENQESLGTETIITWKDFQKTMPEIIVLGGVAGALSGSKS 484
QY 419 GLNIMISKAMQVLGLPNIIVQSIITFVLNPFSAFNANVNVANIVLPILCEMSLALHP 478
DB 485 GLSTWIGNQMLSLSLPWPATLLACILVSVITFEVSNPATITFLPILCSLSETQINP 544
QY 479 LILTLPACLGISMVYFLPVSTPPNAIVTOYAHIKTKYFACCGIVPTIIGISVALVNTW 538
DB 545 LYTILPVTMCLISFVAMLPVGNPPNAIVFSYGHCHQIKDMVKAGLVGNVIGLVIMVAINW 604
QY 539 GLIIFPESKSPDWAKEIKNOT 560
DB 605 GVSLEF-HLDTYPAWAR-VSNIT 624

RESULT 13

US-09-949-016-9977

; Sequence 9977, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9977
; LENGTH: 601
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-9977

Query Match 24.0%; Score 696; DB 4; Length 601;
Best Local Similarity 29.5%; Pred. No. 2.9e-61;
Matches 174; Conservative 121; Mismatches 224; Indels 70; Gaps 14;
QY 27 IIPITLPIIYGFQTDMAEFCGLWLVITMALLMITETLPYVTVLPLVFCPLGLLVN 86
DB 26 VVLLPLPVLH---TKEACAYTLFVVAITFVATFVTEALPLSVLTPSLMPFIMPGIMPS 80
QY 87 SIVCKQFTDIIVFLGLIIVGALGIEYSLHTRIALRVIRVGGSPRRFLVGLMSVTFM 146
DB 81 KKVASAFKDPHLLIGVICLATSIEKWNHKLRLMVLMAKAPGMLLFCMCCTTL 140
QY 147 GLWISNSAGTAMCPYKVALNELDTNKIFPVYMTQ-----BEE 185
DB 141 SMWLSNISTTANVMPVEAVVQI-INAAEVEATQMTYFNGSTNHGLEIDESVNGHEIN 199
QY 186 PVSEGEPPH-----SKITVAFYAG-----TAYASSIGGL 215
DB 200 ERKEKTKPEVPGYNNDTGKISSKVELEKNSGMRTKYRTKKGHVTRKLTCLCIAYSSIGGL 259
QY 216 GTLTGTGNLVRFGIYTERPPTSTVEITFANFMFYSIP-LMVIIVNVLIIAFILITHMGL 274
DB 260 TTIIGTSTNLIFABYFNTRYDPDCRC-INFOGWSFTFSPALIIILLSWINQWLF--LGF 316
QY 275 -FRPNSKTGKIIAEANTNRKLMEDVLQRHIDLGPMSCHEIQMAIAPFMIIVLLITRKPG 333
DB 317 NFKEMFCKGK---TKTVQKACAEVQKQYKGLGPIRYQEIIVTLVLFIMALLWFSRDPG 373
QY 334 FVPGWSDNLNRKVVGS- -SGLSPIVLLIIPALPQYTFYFYCCGCGPFTAQAIADILSWE 391
DB 374 FVPGWSALFS-EYPGFATDSTVALLIGLFLPLIPAKTLTK-TTPTGEIVAFDYSPITWK 431
QY 392 YVLNRNIPWGLLFLGCGFALAVASRETGLNIMISKAMQVLGLPNIIVQSIITFVLNPF 451
DB 432 EFQSFMDTALVGGGFALADGCEESGLSKWNKLSPLGSLPAWLIILLSSLMVTSLT 491
QY 452 AFNANVNVANIVLPILCEMSLALHPILTLPACLGISMVYFLPVSTPPNAIVTOYAH 511
DB 492 EVASNPATITFLPILSPALAEIHNPLXILIPSTLCTSPAFLLPVANPPNAIVFSYGH 551
QY 512 KTKYFACCGIVPTIIGISVALVNTWGLIIFPESKSPDWAKEIKNOT 560
DB 552 KVIDMWKAGLVGNVIGVAVVMLGICTWIVPMF-DLYTYPSPWAPMSNET 599

RESULT 14

US-09-270-767-57987
; Sequence 57987, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:

; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 57987
; LENGTH: 132
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-57987

Query Match 22.8%; Score 661; DB 4; Length 132;
Best Local Similarity 100.0%; Pred. No. 1.1e-58;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 383 AIDAILSWYVLRNIPWGLLFLGGGFALAVASRETGLNIMISKAMQVLIQLPNIVVQSI 442
Db 1 AIDAILSWYVLRNIPWGLLFLGGGFALAVASRETGLNIMISKAMQVLIQLPNIVVQSI 60

Qy 443 TFVLNPFSAFNANVVANIVLPILCEMSLALHPLILTLPAICLISGMVYFLPVSTPPN 502
Db 61 TFVLNPFSAFNANVVANIVLPILCEMSLALHPLILTLPAICLISGMVYFLPVSTPPN 120

Qy 503 AIVTOYAHIKTK 514
Db 121 AIVTOYAHIKTK 132

RESULT 15
US-09-556-916-24
; Sequence 24, Application US/09556916
; Patent No. 6548271
; GENERAL INFORMATION:
; APPLICANT: Turner, Alex
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Sande, Arthur T.
; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins
; FILE REFERENCE: 8535-041-999
; CURRENT APPLICATION NUMBER: US/09/556,916
; CURRENT FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 24
; LENGTH: 580
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-556-916-24

Query Match 21.5%; Score 624; DB 4; Length 580;
Best Local Similarity 25.4%; Pred. No. 5.1e-54;
Matches 158; Conservative 121; Mismatches 196; Indels 146; Gaps 16;

Qy 26 IILPLITLPIIYGFQDIAEFKCLMLIVTMALLWITETLPIYVVTALFPLVFCPLLGLVN 85
Db 17 VCVPELLLPVLV---HPSSEASCAVLIIVTAVYVWSEAVPLGNAALVPAPFLYPPFGVLR 73

Qy 86 ASIVCKQYFDTIIVVFLGGLIYALGIEYSNLHTRIALRVIRIVGGSRRRLFVGLMSVSTF 145
Db 74 SNEVAAYEFKNTLLLVGVICVAAAEKVNHLKRIALRMVLMAGAKPGMLLCLFCMCCTIL 133

Qy 146 MGLWISNSAGTMMCPVIVKALVNL-----DTNKIFPVYMTQEEBPV----- 187
Db 134 LSWLSNTSTTAMVMPVIEAVLQELVSAEDELVAGNSNTEEAEPISLDVKNQPSLELI 193

Qy 188 -----EEGEPHPK----- 197
Db 194 FVNEESNADLTLMNENLNGVPSITNPIKTANQHQKQHPQKQVLTPTSPRKQKLN 253
Qy 198 -----ITVAFYAGIAYASSIGGLTGLTGTLNLFVFRGIYTERPFTSTVEITPANF 247

Db 254 RKYGRSHDQMICKCLSLSSISYSATIGGLTIIGTSTSLIFLHFNNQYPAAEV-VNFGTW 312
Qy 248 MEYSIPLAVIVNVTVIIAFLITHMGLRPNRSK-TGKIIABANTNR-KLMEDVLRORHID 305
Db 313 FLFSFPISLI-----MLVVSFWFMHMLFLGCNFKETCSLSKSKKTKREQLSEKRIQEEYK 368
Qy 306 LGPMSCHETQMAIAFAFMIVLILITRKPGFVPCGWSDLINRKVVGSASGLS-FIVLILFALP 364
Db 369 LGDISYPMVVTGFFFILMTVLFWTREPFGVPCGWDSPFEKKGYRTDATVSVFLGLFLIP 428

Qy 365 TQYTFKYCCGK---GPFTAQAI--DAILSWYVLRNIPWGLLFLGGGFALAVASRETG 419
Db 429 AK----KPCFGKKNDCNGEHSLSGTESIITWKDFOKTMEPWEIVILVGGGYALASGSKSG 484

Qy 420 LNMISKAMQVLIQLPNIVVQSIITFVLNPFSAFNANVVANIVLPILCEMSLALHPL 479
Db 485 LSTWIGNQMLSSLPFPWAVTLLACILVSIVTEFVSNPATITIFLPILCSL----- 535

Qy 480 ILTLPAICLISGMVYFLPVSTPPNAIVTOYAHIKTKYFACCGIVPTIIGISVALVNTNTWG 539
Db 536 ---VKAGLGVN-----VIGLVIVMVAINTWG 558

Qy 540 LIIFPESKSFDPWAKEIKNOT 560
Db 559 VSLF-HLDTYPAWAR-VSNIT 577

Search completed: June 30, 2005, 09:02:13
Job time : 23.608 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2005, 08:59:01 ; Search time 72.9269 Seconds
(without alignments)
2973.073 Million cell updates/sec

Title: US-10-017-479A-3

Perfect score: 2897

Sequence: 1 MAEPGEQRKFVIGRCIFHW.....FPESKSPDPWAKEIKNQTKI 562

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1722976 seqs, 385795295 residues

Total number of hits satisfying chosen parameters: 1722976

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09D_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US10F_PUBCOMB.pep.*
- 19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
- 20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2772	95.7	562	14	US-10-167-994-10
2	1382.5	47.7	572	14	US-10-167-994-2
3	1382.5	47.7	572	14	US-10-017-479-2
4	1382.5	47.7	572	17	US-10-718-359-2
5	885	30.5	568	15	US-10-092-900A-44
6	885	30.5	568	15	US-10-403-161-70
7	875	30.2	568	13	US-10-173-519-2
8	875	30.2	568	15	US-10-403-161-72
9	875	30.2	568	16	US-10-490-080-1
10	875	30.2	568	17	US-10-718-359-6
11	874.5	30.2	568	16	US-10-128-558-149
					Sequence 10, Appl
					Sequence 2, Appl
					Sequence 2, Appl
					Sequence 2, Appl
					Sequence 44, Appl
					Sequence 70, Appl
					Sequence 2, Appl
					Sequence 72, Appl
					Sequence 1, Appl
					Sequence 6, Appl
					Sequence 149, App

871	30.1	587	13	US-10-173-519-5	Sequence 5, Appl
871	30.1	587	14	US-10-167-994-12	Sequence 12, Appl
871	30.1	587	17	US-10-718-359-13	Sequence 13, Appl
868	30.0	616	15	US-10-092-900A-46	Sequence 46, Appl
868	30.0	616	15	US-10-092-900A-48	Sequence 48, Appl
868	30.0	616	15	US-10-403-161-68	Sequence 68, Appl
863	29.8	568	9	US-09-729-094-2	Sequence 2, Appl
863	29.8	568	14	US-10-435-631-2	Sequence 2, Appl
863	29.8	568	14	US-10-167-994-6	Sequence 6, Appl
855.5	29.5	572	17	US-10-718-359-4	Sequence 4, Appl
852.5	29.4	619	9	US-09-729-094-4	Sequence 4, Appl
852.5	29.4	619	14	US-10-435-631-4	Sequence 4, Appl
849.5	29.3	581	17	US-10-718-359-12	Sequence 12, Appl
844	29.1	592	14	US-10-167-994-3	Sequence 3, Appl
844	29.1	592	14	US-10-167-994-11	Sequence 11, Appl
836.5	28.9	539	16	US-10-343-903-25	Sequence 25, Appl
833.5	28.8	602	14	US-10-167-994-4	Sequence 4, Appl
833	28.8	600	14	US-10-167-994-13	Sequence 13, Appl
833	28.8	600	17	US-10-718-359-14	Sequence 14, Appl
814.5	28.1	572	16	US-10-490-080-23	Sequence 23, Appl
814.5	28.1	572	17	US-10-718-359-10	Sequence 10, Appl
743	25.6	626	14	US-10-368-687-20	Sequence 20, Appl
742.5	25.6	627	14	US-10-368-687-8	Sequence 8, Appl
742	25.6	627	15	US-10-092-900A-50	Sequence 50, Appl
742	25.6	627	15	US-10-403-161-74	Sequence 74, Appl
741	25.6	626	14	US-10-368-687-14	Sequence 14, Appl
740.5	25.6	627	14	US-10-368-687-2	Sequence 2, Appl
738.5	25.5	627	14	US-10-167-994-7	Sequence 7, Appl
737	25.4	626	14	US-10-368-687-22	Sequence 22, Appl
736.5	25.4	627	14	US-10-368-687-10	Sequence 10, Appl
735	25.4	626	14	US-10-368-687-16	Sequence 16, Appl
734.5	25.4	627	14	US-10-368-687-4	Sequence 4, Appl
698	24.1	520	15	US-10-108-260A-3862	Sequence 3862, Ap
696	24.0	595	14	US-10-034-934-118	Sequence 118, App
696	24.0	595	14	US-10-167-994-5	Sequence 5, Appl
696	24.0	595	15	US-10-380-727-9	Sequence 9, Appl
693.5	23.9	516	15	US-10-092-900A-52	Sequence 52, Appl
693.5	23.9	516	15	US-10-403-161-76	Sequence 76, Appl
671	23.2	551	17	US-10-718-359-8	Sequence 8, Appl
624	21.5	580	14	US-10-368-687-24	Sequence 24, Appl
623.5	21.5	581	14	US-10-368-687-12	Sequence 12, Appl
622	21.5	580	14	US-10-368-687-18	Sequence 18, Appl
621.5	21.5	581	14	US-10-368-687-6	Sequence 6, Appl
570	19.5	548	17	US-10-437-963-127443	Sequence 127443,
566	19.5	548	17	US-10-928-992-160	Sequence 160, App
551	19.0	510	9	US-09-738-626-3754	Sequence 3754, Ap
549	19.0	548	16	US-10-437-963-127444	Sequence 127444,
545	18.8	552	10	US-09-882-227-496	Sequence 496, App
544	18.8	527	15	US-10-627-476-516	Sequence 516, App
535	18.5	407	11	US-09-833-245-1947	Sequence 1947, Ap
533	18.4	531	16	US-10-425-115-250656	Sequence 250656,
511.5	17.7	377	11	US-09-833-245-1949	Sequence 1949, Ap
487.5	16.8	382	16	US-10-128-558-346	Sequence 346, App
433.5	15.0	446	15	US-10-369-493-21334	Sequence 21334, A
410	14.2	304	15	US-10-264-237-2022	Sequence 2022, Ap
383.5	13.2	376	15	US-10-424-599-254988	Sequence 254988,
351	12.1	233	15	US-10-351-334-176	Sequence 176, App
345.5	11.9	132	13	US-10-173-519-7	Sequence 7, Appl
337.5	11.6	189	15	US-10-276-774-2634	Sequence 2634, Ap
337.5	11.3	335	15	US-10-627-476-518	Sequence 518, App
308	10.6	1006	17	US-10-741-849-7011	Sequence 7011, Ap
306.5	10.6	214	10	US-09-974-879-354	Sequence 354, App
306.5	10.6	214	10	US-09-305-736-354	Sequence 354, App
306.5	10.6	214	11	US-09-818-683-354	Sequence 354, App
306.5	10.6	214	11	US-09-818-683-354	Sequence 354, App
306.5	10.6	214	15	US-10-621-401-354	Sequence 354, App
306.5	10.6	453	15	US-10-369-493-9936	Sequence 9936, Ap
303.5	10.5	585	13	US-10-173-519-4	Sequence 4, Appl
295	10.2	177	13	US-10-173-519-8	Sequence 8, Appl
293.5	9.8	518	17	US-10-470-048B-584	Sequence 584, App
259	8.9	499	14	US-10-238-075-500	Sequence 500, App
252.5	8.7	162	11	US-10-833-245-1948	Sequence 1948, Ap
251.5	8.7	146	15	US-10-296-115-1187	Sequence 1187, Ap

```
85      251      8.7      162      11      US-09-833-245-1945      Sequence 1945, Ap
86      250.5      8.6      236      16      US-10-767-701-43427      Sequence 43427, A
87      224.5      7.7      174      16      US-10-425-115-335649      Sequence 335649,
88      223.5      7.7      226      15      US-10-296-115-980          Sequence 980, App
89      220.5      7.6      300      16      US-10-767-701-46317      Sequence 46317, A
90      214.5      7.4      259      16      US-10-739-930-9458       Sequence 9458, Ap
91      207       7.1      78       9       US-09-864-761-34252      Sequence 34252, A
92      207       7.1      78      14      US-10-029-386-28038      Sequence 28038, A
93      207       7.1      163      15      US-10-424-599-184881     Sequence 184881,
94      206.5      7.1      234      15      US-10-335-977-5397       Sequence 5397, Ap
95      185       6.4      457      15      US-10-369-493-4366       Sequence 4266, Ap
96      180       6.2      429      16      US-10-482-706-191        Sequence 191, App
97      178       6.1      135      15      US-10-276-774-1592       Sequence 1592, Ap
98      169.5      5.9      494      15      US-10-335-977-5882       Sequence 5882, Ap
99      169       5.8      164      15      US-10-424-599-205152     Sequence 205152,
100     165.5      5.7      482      15      US-10-335-977-5881       Sequence 5881, Ap
```

ALIGNMENTS

```
RESULT 1
US-10-167-994-10
; Sequence 10, Application US/10167994
; Publication No. US20030082647A1
; GENERAL INFORMATION:
; APPLICANT: Reenan, Robert A.
; APPLICANT: Rogina, Blanka
; APPLICANT: Helfand, Stephen L.
; TITLE OF INVENTION: TRANSPORTER PROTEIN
; FILE REFERENCE: 13407-013001
; CURRENT APPLICATION NUMBER: US/10/167,994
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/255,013
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 10/017,479
; PRIOR FILING DATE: 2001-12-12
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-167-994-10
```

```
Query Match      95.7%; Score 2772; DB 14; Length 562;
Best Local Similarity 92.0%; Pred. No. 5.4e-250;
Matches 519; Conservative 39; Mismatches 2; Indels 4; Gaps 2;

Qy      1      MAEPGEQRKFFVLGRCCIFHWKASIIPLITLPIIYGFQTDMAEFKCLWLIVTMALLW 60
Db      1      MAEPGEQRKFFVLGRCCIFHWKASIIPLITLPIIYGFQTDMAEFKCLWLIVTMALLW 60

Qy      61      ITETLPIYVTALFPLVFCPLGLVNASIVCKQYFTDTIVVFLGGLI VALGIEYSNLHTRI 120
Db      61      L TETLPIYVTALFPLVFCPLGLVNASIVCKQYFTDTIVVFLGGLI VALGIEYSNLHTRI 120

Qy      121      ALRVIRIVGGSPRRLLFVGLMSVTFMGLWINSAGTAMWCPIVKALVNELDTNKLFPVYM 180
Db      121      ALRVIRIVGGSPRRLLFVGLMSVTFMGLWLSNSAGTAMVCPVVKALVNELDTNKLFPVYM 180

Qy      181      TOEEEPVEEGPPHPSKITVAFYAGIAYASSIGGLTGLTGTLNLFVFGIYTERFPTSTV 240
Db      181      TOEEEPVEEGPPHPSKITVAFYAGIAYASSIGGLTGLTGTLNLFVFGIYTERFPTTV 240

Qy      241      EITFANFMYSIPLMVIYNVTLLVIAFLITHMGLFRPNSKTGKIIAEANTNRKLMEDVLR 300
Db      241      EITFANFMYSNPLMVIYNITVLLAFLITM--FRPNSKTGKIIAEANTNRKLMEDVLR 298

Qy      301      QRHDLGPMSCHEIQMATAFAPMIVLLITRKPFGVGHSDLINRKVGSASGLSFIVLLI 360
Db      299      QRHDLGPMSCHEIQMALAFAPMIVLLITRKPFGVGHSDLINRKVGSASGLSFIVLLL 358
```

```
Qy      361      FALPTQYTFYFKYCC--GKGPFYTAQADAILSWEYVLRNIPWGLLELLGGGFALAVASRET 418
Db      359      FALPTQYTFYFKYCCGKGGKGPYTAQADAILSWEYVLRNIPWGLLELLGGGFALAVASRET 418

Qy      419      GUNIMISKAMQVILGELPMIVQSIITFVLANPFSANVNVVANI VLPILCENSLELHP 478
Db      419      GUNLMLSKAMQVILGLPLVQSIITFVLANPFSANVNVVANI VLPILCENSLELHP 478

Qy      479      LTLTPACILGISMVYFLVPSTPPNAIVTQYAHIKTKYFACCGIVPTIIGISVALVNTNTW 538
Db      479      LLLTLPACILGISMVYFLVPSTPPNAIVTQYAHLKTKYFACCGIVPTLLGLSVALVNTNTW 538

Qy      539      GLIIPESKSFDPDAKEIKNQTKI 562
Db      539      GLIIPESKSFDPDAKELKNQTKL 562
```

```
RESULT 2
US-10-167-994-2
; Sequence 2, Application US/10167994
; Publication No. US20030082647A1
; GENERAL INFORMATION:
; APPLICANT: Reenan, Robert A.
; APPLICANT: Rogina, Blanka
; APPLICANT: Helfand, Stephen L.
; TITLE OF INVENTION: TRANSPORTER PROTEIN
; FILE REFERENCE: 13407-013001
; CURRENT APPLICATION NUMBER: US/10/167,994
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/255,013
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 10/017,479
; PRIOR FILING DATE: 2001-12-12
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 572
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-167-994-2

Query Match      47.7%; Score 1382.5; DB 14; Length 572;
Best Local Similarity 46.5%; Pred. No. 4.2e-120;
Matches 262; Conservative 120; Mismatches 164; Indels 17; Gaps 7;

Qy      3      EPGEQRKFFVLGRCCIF---HWKASIIPLITLPIIYGFQTDMAEFKCLWLIVTMALL 59
Db      4      ETGEQPPPV-KCSNFFANHWKGLVPLVPLCLPVMLL--NEGAEFRCMYLLVMAIF 59

Qy      60      WITETLPIYVTALFPLVFCPLGLVNASIVCKQYFTDTIVVFLGGLI VALGIEYSNLHTR 119
Db      60      WYTEALPLVYTSIMPIVAFPIMGINSDDQTCRLYFKDITLVFMFGGIMVALAVEYCNLHKR 119

Qy      120      IALRVIRIVGGSPRRLLFVGLMSVTFMGLWINSAGTAMWCPIVKALVNELDTNKLFPVY 179
Db      120      LALRVIRIVGGSPRRLLFVGLIMVTFMGLWINSAGTAMWCPIIQAIVLELQAGVCKIN 179

Qy      180      MTQEEPPV-----EAGEPPHPSKITVAFYAGIAYASSIGGLTGLTGTLNLFVFGIYTE 233
Db      180      HEQYQIVGVGNKKNNEDEPPYPTKITLYCYLGIAYASSIGGCGTIIGTATNLAFKGIYE 239

Qy      234      RPTSTVEITFANFMFYSIPLMVINVTLLVIAFLITH--MGLFRPNSKTGKIIAEANTNR 292
Db      240      RFKNSTEQMDPFTFMFYSVPSMLV--YTLTTFVFLQWFMGLWRPKSKEAQVQREGA 297

Qy      293      KLMEVLRORHIDLGPMSCHIEIQMATAFAPMIVLLITRKPFGVGHSDLINRKVGSASG 352
Db      298      DVAKKVIDQRYKOLGPMSIHEIQMILFI FVNMVYFRKFGIFLGWADLLNSKDIRNSMP 357

Qy      353      LSFIVLLI FALPTQYTFYFKYCCGK--GPFYTAQADAILSWEYVLRNIPWGLLELLGGGFAL 411
Db      358      TIFVVMCFMLPANYAFRLYCTRGGPVPTGPTPSLITWKFIQTKVPWGLVFLILGGGFAL 417
```

QY 412 AVASRETLNIMISKAMQVLIGLPIVQSVITFVLNFFSAFNAVNVVNIPLILCEMS 471
Db 418 AEGSKQSMAGLIGNALIGLKLPNLSVLLVTLVAVFLTSSNVAIANIIPVLAEMS 477
QY 472 LALEHPLILTPACLGISVMVFLPSTPPNAIVTOYAHKTKYFACCGIVPTIIGISVA 531
Db 478 LAIEHPLVLLPAGLACSMFHLFVSTPPNALVAGYANIRTKDMAIAGIGPTIITL 537
QY 532 LVNTNTWGLIIPESKSPFDWAK 554
Db 538 FVFCQTWGLVVPNLNSFPPEWAQ 560

RESULT 3

US-10-017-479-2
; Sequence 2, Application US/10017479
; Publication No. US20030104399A1
; GENERAL INFORMATION:
; APPLICANT: University of Connecticut
; APPLICANT: Helfand, Stephan L
; APPLICANT: Reenan, Robert A
; APPLICANT: Rogina, Blanka
; TITLE OF INVENTION: Polynucleotides Encoding Cellular Transporters and Methods of Use
; FILE REFERENCE: UCT-0020
; CURRENT APPLICATION NUMBER: US/10/017,479
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/255,013
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 572
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-017-479-2

Query Match 47.7%; Score 1382.5; DB 14; Length 572;
Best Local Similarity 46.5%; Pred. No. 4.2e-120;
Matches 262; Conservative 120; Mismatches 164; Indels 17; Gaps 7;
QY 3 EPGEQKFLVGRCCIF---HWRGKASIIPLITLPIIYGFQTDMAEKFCLWLVTMALL 59
Db 4 EIGEQQPPV-KCSNFFANHWKGLVFLVPLCLPVMLL---NEGAERFCMYLLLVMAIF 59
QY 60 WITETLPIVYVLTALPFLVPCLLGLVNASIVCKQVFTDIIVVFLGGLIIVAGLIEYNSLHTR 119
Db 60 WYTEALPLYVTSMTPIVAFPIGIMSSDOTCLYFKDTLVWFMGGMVALAVEYCNLHTR 119
QY 120 IALRVIRIVGSPRLFLVGLMSVTFMGLWISNSAGTAMMCPIVKALVNELDTNKIFPVY 179
Db 120 LALRVIOIVGCSPRRLHFLGIMVTFMGLWISNSAGTAMMCPIIQAVLEELQAQGVCKIN 179
QY 180 MTQEEEPV-----EGBEPHPKTIIVAFYAGIYASSIGLGTIGTNLVPRGIYTE 233
Db 180 HEPOYQIVGGNKNNEDEPPYTKITLCYLGIAYASSIGCGGTIGTATNLTKFIYE 239
QY 234 RPTSTVEITFANPMFYISIPLMVIVNVTLVIAFLITH-MGLFRPNSKTGKIIAEANTNR 292
Db 240 RFKSTQMDPTFMFYSVPSMLV--YTLTFVFLQWFMGLWRPKSKAEQVQRGREGA 297
QY 293 KLMDVLQRHIDLGPMSCHEIQMAIAFAMVILLITRKPGFVPGWSDLNIRKVVGSAG 352
Db 298 DVAKKVIDQRYKDLGPMSEIHEIQVMILFIMVVMYFTRKPGIFLGWADLLNSKDIRNSMP 357
QY 353 LSFIVLLIFALPTQYTFPKYCCGK-GPETAQIDAISLWSEYVLRNIPWGLIFLGGG 411
Db 358 TIFVVMCFMPLPANYAFURYCTRRGGVPVTPGTPTSLITWKFIQTKVPMGLVFLGGG 417
QY 412 AVASRETLNIMISKAMQVLIGLPIVQSVITFVLNFFSAFNAVNVVNIPLILCEMS 471
Db 418 AEGSKQSMAGLIGNALIGLKLPNLSVLLVTLVAVFLTSSNVAIANIIPVLAEMS 477
QY 472 LALEHPLILTPACLGISVMVFLPSTPPNAIVTOYAHKTKYFACCGIVPTIIGISVA 531
Db 478 LAIEHPLVLLPAGLACSMFHLFVSTPPNALVAGYANIRTKDMAIAGIGPTIITL 537
QY 532 LVNTNTWGLIIPESKSPFDWAK 554
Db 538 FVFCQTWGLVVPNLNSFPPEWAQ 560

Db 478 LAIEHPLVLLPAGLACSMFHLFVSTPPNALVAGYANIRTKDMAIAGIGPTIITL 537
QY 532 LVNTNTWGLIIPESKSPFDWAK 554
Db 538 FVFCQTWGLVVPNLNSFPPEWAQ 560
RESULT 4
US-10-718-359-2
; Sequence 2, Application US/10718359
; Publication No. US20050095240A1
; GENERAL INFORMATION:
; APPLICANT: MEDICAL COLLEGE OF GEORGIA RESEARCH INSTITUTE, INC.
; TITLE OF INVENTION: NACT AS A TARGET FOR LIFESPAN EXPANSION AND WEIGHT REDUCTION
; FILE REFERENCE: 275,0008,0101
; CURRENT APPLICATION NUMBER: US/10/718,359
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/428,469
; PRIOR FILING DATE: 2002-11-22
; PRIOR APPLICATION NUMBER: 60/459,441
; PRIOR FILING DATE: 2003-04-01
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 572
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-718-359-2

Query Match 47.7%; Score 1382.5; DB 17; Length 572;
Best Local Similarity 46.5%; Pred. No. 4.2e-120;
Matches 262; Conservative 120; Mismatches 164; Indels 17; Gaps 7;
QY 3 EPGEQKFLVGRCCIF---HWRGKASIIPLITLPIIYGFQTDMAEKFCLWLVTMALL 59
Db 4 EIGEQQPPV-KCSNFFANHWKGLVFLVPLCLPVMLL---NEGAERFCMYLLLVMAIF 59
QY 60 WITETLPIVYVLTALPFLVPCLLGLVNASIVCKQVFTDIIVVFLGGLIIVAGLIEYNSLHTR 119
Db 60 WYTEALPLYVTSMTPIVAFPIGIMSSDOTCLYFKDTLVWFMGGMVALAVEYCNLHTR 119
QY 120 IALRVIRIVGSPRLFLVGLMSVTFMGLWISNSAGTAMMCPIVKALVNELDTNKIFPVY 179
Db 120 LALRVIOIVGCSPRRLHFLGIMVTFMGLWISNSAGTAMMCPIIQAVLEELQAQGVCKIN 179
QY 180 MTQEEEPV-----EGBEPHPKTIIVAFYAGIYASSIGLGTIGTNLVPRGIYTE 233
Db 180 HEPOYQIVGGNKNNEDEPPYTKITLCYLGIAYASSIGCGGTIGTATNLTKFIYE 239
QY 234 RPTSTVEITFANPMFYISIPLMVIVNVTLVIAFLITH-MGLFRPNSKTGKIIAEANTNR 292
Db 240 RFKSTQMDPTFMFYSVPSMLV--YTLTFVFLQWFMGLWRPKSKAEQVQRGREGA 297
QY 293 KLMDVLQRHIDLGPMSCHEIQMAIAFAMVILLITRKPGFVPGWSDLNIRKVVGSAG 352
Db 298 DVAKKVIDQRYKDLGPMSEIHEIQVMILFIMVVMYFTRKPGIFLGWADLLNSKDIRNSMP 357
QY 353 LSFIVLLIFALPTQYTFPKYCCGK-GPETAQIDAISLWSEYVLRNIPWGLIFLGGG 411
Db 358 TIFVVMCFMPLPANYAFURYCTRRGGVPVTPGTPTSLITWKFIQTKVPMGLVFLGGG 417
QY 412 AVASRETLNIMISKAMQVLIGLPIVQSVITFVLNFFSAFNAVNVVNIPLILCEMS 471
Db 418 AEGSKQSMAGLIGNALIGLKLPNLSVLLVTLVAVFLTSSNVAIANIIPVLAEMS 477
QY 472 LALEHPLILTPACLGISVMVFLPSTPPNAIVTOYAHKTKYFACCGIVPTIIGISVA 531
Db 478 LAIEHPLVLLPAGLACSMFHLFVSTPPNALVAGYANIRTKDMAIAGIGPTIITL 537
QY 532 LVNTNTWGLIIPESKSPFDWAK 554
Db 538 FVFCQTWGLVVPNLNSFPPEWAQ 560


```
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 70
; LENGTH: 568
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-403-161-70

Query Match      30.5%; Score 885; DB 15; Length 568;
Best Local Similarity 35.6%; Pred. No. 1.3e-73;
Matches 197; Conservative 101; Mismatches 219; Indels 36; Gaps 9;

QY 27 IIPLTLPILYVGFQTDMAEFKCLMLVITMALLMTETLPIYVYVTFPLVFCPLGLVNA 86
DB 19 VTPLLLPLVI-----LMPAKFVRCAYVILMAIYVCTEVIPLAVTSLMPVLLPFLQILDS 74
QY 87 SVVCQYQYFDTTIVFLGGLIIVAGIEYSNLHTRIALRVIRVGGSPRRFLVGLMSVTFM 146
DB 75 ROVCVQYMKDNTMLFLGGLIIVAVAVERNLHRIALRTLWVGAKPARMLGFMGVTTALL 134
QY 147 GLWISNAGTAMMCPITVKALVNLDTN-----KIFPVYMTQEEEP-VVE 189
DB 135 SMWISNTATTAMVPIVEAILQOMEATSAATEAGLELDVKGKAKELPGSQVIFEGPTLQ 194
QY 190 GEPHPHSKITVAFYAGIAYASSIGGLTGLTGTLNLFVFRGIYTERPPTSTVETITFANMF 249
DB 195 QEDQERKRLCKAMTLCICYAASIGGTATLTGTGPNVVLGQMNELFPDSKOLVNFASWFA 254
QY 250 YSIPLMVIVNVLVIAEL-ITHMGLFRNSKTGKIIABANTNRKLMEDVLRQRIHIDLP 308
DB 255 FAFPNMLV-----MLLFAWMLQFYVYKRFNFKKSWGCGLESKKKKAALKVLOEYRKLGP 310
QY 309 MSCHIEIQMAIAFAFMIVLLITRKPFGVPGWSDLI-----NRKVVGSASGLSFIVLLIFALP 364
DB 311 LSFANINVLICFLLVILWFSRDPGFMGMLTVANVEGETKYVSDATVAIFVATLLFIVP 370
QY 365 QYTFEYKVC-----GKGPTTAQIDAILSWEYVLRNIPMGLLFLGGGFALAVASRETL 420
DB 371 SQKPKFNRFSQTEEGKSPVLI-APPPLDWKVQEKVPMGIVLLGGGFALAKGSEASGL 429
QY 421 NIMISKAMQVLGLPNIVVQSTTFVLNFFSFANVNVVAVNIVLILCEMSLALHLP 480
DB 430 SVMGMGQMEPLHVPAAITLTLSLVAVFTECTSNVATTLFLPFIASMSRSIGLNFLY 489
QY 481 LTLPACLGISVMYFLPVPSTPPNAIVTQYAHIKTKYFACCGIVPTIIGISVALVNTWGL 540
DB 490 IMLPCTLSASFAMLPVATPPNAIVFTYGHKLVADMVKTGVIMNIIGVFCVFLAVNTWGR 549
QY 541 IIPESKSPDDWA 553
DB 550 AIF-DLDHFPDWA 561

RESULT 7
US-10-173-519-2
; Sequence 2, Application US/10173519
; Publication No. US20020193582A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Roly A.J.
; APPLICANT: Millennium Pharmaceuticals, Inc.
; TITLE OF INVENTION: 69624, A Human Transporter Family Member
; TITLE OF INVENTION: and Uses Therefor
; FILE REFERENCE: MP101-098P1RM
; CURRENT APPLICATION NUMBER: US/10/173,519
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 60/298,970
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 568
; TYPE: PRT
; ORGANISM: Homo sapiens

US-10-173-519-2
Query Match      30.2%; Score 875; DB 13; Length 568;
Best Local Similarity 35.7%; Pred. No. 1.1e-72;
Matches 198; Conservative 99; Mismatches 218; Indels 40; Gaps 10;

QY 27 IIPLTLPILYVGFQTDMAEF-KCLMLVITMALLMTETLPIYVYVTFPLVFCPLGLVNA 85
DB 19 VTPLLLPLVI-----LMPAKFVRCAYVILMAIYVCTEVIPLAVTSLMPVLLPFLQILDS 74
QY 86 ASIYCKQYFDTTIVFLGGLIIVAGIEYSNLHTRIALRVIRVGGSPRRFLVGLMSVTF 145
DB 75 SRQVCVQYMKDNTMLFLGGLIIVAVAVERNLHRIALRTLWVGAKPARMLGFMGVTTALL 134
QY 146 MGLWISNAGTAMMCPITVKALVNLDTN-----KIFPVYMTQEEEP-VVE 188
DB 135 LSWISNTATTAMVPIVEAILQOMEATSAATEAGLELDVKGKAKELPGSQVIFEGPTLQ 194
QY 189 GEPHPHSKITVAFYAGIAYASSIGGLTGLTGTLNLFVFRGIYTERPPTSTVETITFANMF 248
DB 195 QEDQERKRLCKAMTLCICYAASIGGTATLTGTGPNVVLGQMNELFPDSKOLVNFASW 254
QY 249 FYSIPLMVIVNVLVIAEL-ITHMGLFRPN-SKTGKIIABANTNRKLMEDVLRQRIHID 307
DB 255 AFAFPNMLV-----MLLFAWMLQFYVYKRFNFKKSWGCGLESKKKKAALKVLOEYRKL 310
QY 308 PMSCHIEIQMAIAFAFMIVLLITRKPFGVPGWSDLI-----NRKVVGSASGLSFIVLLIFALP 363
DB 311 PLSFAINVLICFLLVILWFSRDPGFMGMLTVANVEGETKYVSDATVAIFVATLLFIV 370
QY 364 PTQYTFEYKVC-----GKGPTTAQIDAILSWEYVLRNIPMGLLFLGGGFALAVASRET 418
DB 371 PSQKPKFNRFSQTEEGKTPFPV---PELLOWKVQEKVPMGIVLLGGGFALAKGSEAS 427
QY 419 GLNIMISKAMQVLGLPNIVVQSTTFVLNFFSFANVNVVAVNIVLILCEMSLALHLP 478
DB 428 GLSVMGMGQMEPLHVPAAITLTLSLVAVFTECTSNVATTLFLPFIASMSRSIGLN 487
QY 479 LILTLPACLGISVMYFLPVPSTPPNAIVTQYAHIKTKYFACCGIVPTIIGISVALVNTW 538
DB 488 LYIMLPCTLSASFAMLPVATPPNAIVFTYGHKLVADMVKTGVIMNIIGVFCVFLAVNTW 547
QY 539 GLIIPESKSPDDWA 553
DB 548 GRAIF-DLDHFPDWA 561

RESULT 8
US-10-403-161-72
; Sequence 72, Application US/10403161
; Publication No. US20040043930A1
; GENERAL INFORMATION:
; APPLICANT: Angerson, David et al.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-573C
; CURRENT APPLICATION NUMBER: US/10/403,161
; CURRENT FILING DATE: 2003-03-31
; PRIOR APPLICATION NUMBER: 60/370349
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/384543
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/370969
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 60/403748
; PRIOR FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: 60/372019
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/374379
; PRIOR FILING DATE: 2002-04-22
; PRIOR APPLICATION NUMBER: 05/779679
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/181045
; PRIOR FILING DATE: 2000-02-08
```

; PRIOR APPLICATION NUMBER: 10/055877
; PRIOR FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: 60/262892
; PRIOR FILING DATE: 2001-01-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 72
; LENGTH: 568
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-403-161-72

Query Match 30.2%; Score 875; DB 15; Length 568;
Best Local Similarity 35.7%; Pred. No. 1.1e-72;
Matches 198; Conservative 99; Mismatches 218; Indels 40; Gaps 10;
QY 27 IIPILITPILYIGFQTDMAEF-KCLWLIVTMALLMITETPLPIYVATLPLVFCPLGLV 85
DB 19 VTPLLLLPLVI-----LMPAKFVRCAYVILMAIYWCTEVIPLAVTSLMPVLLPFLQILD 74
QY 86 ASIVCKQYFTDTIVVFLGGLIVALGIEYSNLHTRIALRVIRVGGSPRLFLVGLMSVSIF 145
DB 75 SRQVCVQYMKDTNMLFLGGLIIVAVAVERNLHRIALRTLLWVGAKPARMLGFMGV 134
QY 146 MGLWISNSAGTAMMCPIVKALVNELDN-----KIFPVYMTQEEEP-VE 188
DB 135 LSMWISNTATTAMVPIVEAILQOMEATSAATEAGLELVKDGKAKELPGSQVIFEGTILG 194
QY 189 EGEPHPKSKITVAFYAGIAYASSIGGLTIGTGNLVFRGIYTERPPTSTVITFANFM 248
DB 195 QOEQDERKRLCKAMTLCICVAAASIGGTATLTGTGPNVLLGQNNELFPDSDKLVNFASF 254
QY 249 FYSIPLMVIVNVLIIAFLITHMGLFRPN-SKTGKIIAEANTNRKLMEDVLRQRHIDL 307
DB 255 AFAPFNMLV-----MLLFAWLMLQFVYMRNFKKSMGCGLESKKNEAKALKVLOEYRKLG 310
QY 308 PMSCHIEOMAIAPAFMIVLLITRKPGFVPGWSDLI-----NRKVGSASGLSFTVLILFAL 363
DB 311 PLSPAEINVLICFLLVILWFSRDPGFMGMLTVAVVEGETKYVSDATVAIFVATLLFIV 370
QY 364 PTQYTFKYKCC-----GKGPFTAQAIADAILSWEYVLRNIPWGLLFLGGGFALAVASRET 418
DB 371 PSQPKFNFRSQTEERKTPFP---PPLLDWKVTQEKVPWGIIVLLGGGFALAKGSEAS 427
QY 419 GLNIMISKAMQVLIGLNPVQSIPTFVLANFFSAFNANVAVNIPLILCEMSLALHLP 478
DB 428 GLSVMGKQMEPLHVPVPPAAITLILSLVAVFTECTSNVATTTFLPFIASMSKSI 487
QY 479 LILTLPACLGISMVYFLPVSTPPNAIVTQYAHIKTKYFACCGIVPTIIGISVALVNTNTW 538
DB 488 LYIMLPCTLSAFAMLPVATPNAIVFTYGHKLVADVMKVTGVMIMNIIGVFCVFLAVNTW 547
QY 539 GLIIFPESKSPDDWA 553
DB 548 GRAIF-DLDHFPDWA 561

RESULT 9
US-10-490-080-1
; Sequence 1, Application US/10490080
; Publication No. US20040253597A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.
; TITLE OF INVENTION: Novel Protein and its DNA
; FILE REFERENCE: P02-0109PCT
; CURRENT APPLICATION NUMBER: US/10/490,080
; CURRENT FILING DATE: 2004-03-17
; PRIOR APPLICATION NUMBER: JP 2001-281992
; PRIOR FILING DATE: 2001-09-17
; PRIOR APPLICATION NUMBER: JP 2001-306873
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: JP 2002-113279

; PRIOR FILING DATE: 2002-04-16
; NUMBER OF SEQ ID NOS: 42
; SEQ ID NO 1
; LENGTH: 568
; TYPE: PRT
; ORGANISM: Human
US-10-490-080-1

Query Match 30.2%; Score 875; DB 16; Length 568;
Best Local Similarity 35.7%; Pred. No. 1.1e-72;
Matches 198; Conservative 99; Mismatches 218; Indels 40; Gaps 10;
QY 27 IIPILITPILYIGFQTDMAEF-KCLWLIVTMALLMITETPLPIYVATLPLVFCPLGLV 85
DB 19 VTPLLLLPLVI-----LMPAKFVRCAYVILMAIYWCTEVIPLAVTSLMBVLLPFLQILD 74
QY 86 ASIVCKQYFTDTIVVFLGGLIVALGIEYSNLHTRIALRVIRVGGSPRLFLVGLMSVSIF 145
DB 75 SRQVCVQYMKDTNMLFLGGLIIVAVAVERNLHRIALRTLLWVGAKPARMLGFMGV 134
QY 146 MGLWISNSAGTAMMCPIVKALVNELDN-----KIFPVYMTQEEEP-VE 188
DB 135 LSMWISNTATTAMVPIVEAILQOMEATSAATEAGLELVKDGKAKELPGSQVIFEGTILG 194
QY 189 EGEPHPKSKITVAFYAGIAYASSIGGLTIGTGNLVFRGIYTERPPTSTVITFANFM 248
DB 195 QOEQDERKRLCKAMTLCICVAAASIGGTATLTGTGPNVLLGQNNELFPDSDKLVNFASF 254
QY 249 FYSIPLMVIVNVLIIAFLITHMGLFRPN-SKTGKIIAEANTNRKLMEDVLRQRHIDL 307
DB 255 AFAPFNMLV-----MLLFAWLMLQFVYMRNFKKSMGCGLESKKNEAKALKVLOEYRKLG 310
QY 308 PMSCHIEOMAIAPAFMIVLLITRKPGFVPGWSDLI-----NRKVGSASGLSFTVLILFAL 363
DB 311 PLSPAEINVLICFLLVILWFSRDPGFMGMLTVAVVEGETKYVSDATVAIFVATLLFIV 370
QY 364 PTQYTFKYKCC-----GKGPFTAQAIADAILSWEYVLRNIPWGLLFLGGGFALAVASRET 418
DB 371 PSQPKFNFRSQTEERKTPFP---PPLLDWKVTQEKVPWGIIVLLGGGFALAKGSEAS 427
QY 419 GLNIMISKAMQVLIGLNPVQSIPTFVLANFFSAFNANVAVNIPLILCEMSLALHLP 478
DB 428 GLSVMGKQMEPLHVPVPPAAITLILSLVAVFTECTSNVATTTFLPFIASMSKSI 487
QY 479 LILTLPACLGISMVYFLPVSTPPNAIVTQYAHIKTKYFACCGIVPTIIGISVALVNTNTW 538
DB 488 LYIMLPCTLSAFAMLPVATPNAIVFTYGHKLVADVMKVTGVMIMNIIGVFCVFLAVNTW 547
QY 539 GLIIFPESKSPDDWA 553
DB 548 GRAIF-DLDHFPDWA 561

RESULT 10
US-10-718-359-6
; Sequence 6, Application US/10718359
; Publication No. US20050095240A1
; GENERAL INFORMATION:
; APPLICANT: MEDICAL COLLEGE OF GEORGIA RESEARCH INSTITUTE, INC.
; TITLE OF INVENTION: NaCT AS A TARGET FOR LIFESPAN EXPANSION AND WEIGHT REDUCTION
; FILE REFERENCE: 275.0008 0101
; CURRENT APPLICATION NUMBER: US/10/718,359
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/428,469
; PRIOR FILING DATE: 2002-11-22
; PRIOR APPLICATION NUMBER: 60/459,441
; PRIOR FILING DATE: 2003-04-01
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 568
; TYPE: PRT
; ORGANISM: human NaCT

; CURRENT APPLICATION NUMBER: US/10/173,519
 ; CURRENT FILING DATE: 2002-06-17
 ; PRIOR APPLICATION NUMBER: 60/298,970
 ; PRIOR FILING DATE: 2001-06-18
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 5
 ; LENGTH: 587
 ; TYPE: PRT
 ; ORGANISM: Rattus norvegicus
 US-10-173-519-5

Query Match 30.1%; Score 871; DB 13; Length 587;
 Best Local Similarity 37.4%; Pred. No. 2.8e-72;
 Matches 213; Conservative 101; Mismatches 196; Indels 60; Gaps 17;
 QY 30 LITPLIYGFQTDMAEFKCLMIVTMALLWITETLPIYVTALFPLVFCPLGLGNASIV 89
 Db 24 LLPLPLIV--QT--KEAYCAYSIIIMALLWCTEALPLAVTALFPIVLPFLMGIMDAEV 78
 QY 90 CKQYFTDTIVVFLGGLIVALGIEYSNLHTRIALRVIRIVGSSPRRLFVGLMSVSTFWGLW 149
 Db 79 CIEYFKDTNILFVGLLVAIAVEHNLHRIALQVLLIIGVRPALLLGLFMLVTAFLSMW 138
 QY 150 INSAGTAMCPIVKALVNELDTNK-----IFPVMTQEEBPVEE-----GEP----- 192
 Db 139 ISNTATTAMVPICHAVLEQSGKDVGGNNPTFELQECPOKEVTKLDNGQPVSA 198
 QY 193 --PHPSKITV--AFYAG-----IAYASSIGGLTIGTGNLVRGIVYTRFPSTVEITP 244
 Db 199 SEPRTOQTQEHHRFSQGLSLCICYSASIGGIATLTGTTNVLQGVNSLFPQNGVNVF 258
 QY 245 ANPMFYSIPLMIVNVTLLVIAEL---ITHMGL--FRNSKTGKIIAEANTNRKLM--DVL 299
 Db 259 ASWFGFAFPTMII---LLLAWLQVLFVGNFRKNFGF---EGEERKQAAAFQVI 310
 QY 300 RQRHIDLGPMSCHEIOAIAFAFMIVLLITRKPGFVPGWSDLI--NRKVVGASG---LS 354
 Db 311 KTQYRLLGPMSPAEKTVTVLVLLVLMVTRFPGFPGWGDVTFANEKQSMASDGTVAI 370
 QY 355 FIVLLIFALPTQYTFKYCCG-----KGPFTAQDAIDAILSWEVYVLRNIPWGLLFLGGGF 409
 Db 371 FISLVMEIIPSKIP-----GLMQDPKPKGLKAPPAITLTKTNDKMPWNIIVILLGGGF 424
 QY 410 ALAVASRETGNIMISKAMQVILGLPNIVQSIITFVLNFFSAFNANVVANVILPTLCE 469
 Db 425 ALAKGSEQSGLEWLGDKLTPLQHIPPSATAVILCLLIAIFTECTSNVATTTFLPLILAS 484
 QY 470 MSLALELHPLILTLPACLGISWYFVLPVSTPPNAIVTQYAHIKTKYFACCGIVPTIIGIS 529
 Db 485 MAQAICLHPLYVMLPCTLAASLAFMLPVATPPNAIVFSGGLKVSMDMARAGFLNIIIGVL 544
 QY 530 VALVNTNTWGLIIPPEKSPDPWAKEIKNQ 559
 Db 545 AITLSINSWSIPIF-KLDTFPPSWAHSNTSQ 573

RESULT 13
 US-10-167-994-12
 ; Sequence 12, Application US/10167994
 ; Publication No. US20030082647A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Reenan, Robert A.
 ; APPLICANT: Rogina, Blanka
 ; APPLICANT: Helfand, Stephen L.
 ; TITLE OF INVENTION: TRANSPORTER PROTEIN
 ; FILE REFERENCE: 13407-013001
 ; CURRENT APPLICATION NUMBER: US/10/167,994
 ; CURRENT FILING DATE: 2002-06-12
 ; PRIOR APPLICATION NUMBER: US 60/255,013
 ; PRIOR FILING DATE: 2000-12-12
 ; PRIOR APPLICATION NUMBER: US 10/017,479
 ; PRIOR FILING DATE: 2001-12-12

; NUMBER OF SEQ ID NOS: 13
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 12
 ; LENGTH: 587
 ; TYPE: PRT
 ; ORGANISM: Rattus norvegicus
 US-10-167-994-12
 Query Match 30.1%; Score 871; DB 14; Length 587;
 Best Local Similarity 37.4%; Pred. No. 2.8e-72;
 Matches 213; Conservative 101; Mismatches 196; Indels 60; Gaps 17;
 QY 30 LITPLIYGFQTDMAEFKCLMIVTMALLWITETLPIYVTALFPLVFCPLGLGNASIV 89
 Db 24 LLPLPLIV--QT--KEAYCAYSIIIMALLWCTEALPLAVTALFPIVLPFLMGIMDAEV 78
 QY 90 CKQYFTDTIVVFLGGLIVALGIEYSNLHTRIALRVIRIVGSSPRRLFVGLMSVSTFWGLW 149
 Db 79 CIEYFKDTNILFVGLLVAIAVEHNLHRIALQVLLIIGVRPALLLGLFMLVTAFLSMW 138
 QY 150 INSAGTAMCPIVKALVNELDTNK-----IFPVMTQEEBPVEE-----GEP----- 192
 Db 139 ISNTATTAMVPICHAVLEQSGKDVGGNNPTFELQECPOKEVTKLDNGQPVSA 198
 QY 193 --PHPSKITV--AFYAG-----IAYASSIGGLTIGTGNLVRGIVYTRFPSTVEITP 244
 Db 199 SEPRTOQTQEHHRFSQGLSLCICYSASIGGIATLTGTTNVLQGVNSLFPQNGVNVF 258
 QY 245 ANPMFYSIPLMIVNVTLLVIAEL---ITHMGL--FRNSKTGKIIAEANTNRKLM--DVL 299
 Db 259 ASWFGFAFPTMII---LLLAWLQVLFVGNFRKNFGF---EGEERKQAAAFQVI 310
 QY 300 RQRHIDLGPMSCHEIOAIAFAFMIVLLITRKPGFVPGWSDLI--NRKVVGASG---LS 354
 Db 311 KTQYRLLGPMSPAEKTVTVLVLLVLMVTRFPGFPGWGDVTFANEKQSMASDGTVAI 370
 QY 355 FIVLLIFALPTQYTFKYCCG-----KGPFTAQDAIDAILSWEVYVLRNIPWGLLFLGGGF 409
 Db 371 FISLVMEIIPSKIP-----GLMQDPKPKGLKAPPAITLTKTNDKMPWNIIVILLGGGF 424
 QY 410 ALAVASRETGNIMISKAMQVILGLPNIVQSIITFVLNFFSAFNANVVANVILPTLCE 469
 Db 425 ALAKGSEQSGLEWLGDKLTPLQHIPPSATAVILCLLIAIFTECTSNVATTTFLPLILAS 484
 QY 470 MSLALELHPLILTLPACLGISWYFVLPVSTPPNAIVTQYAHIKTKYFACCGIVPTIIGIS 529
 Db 485 MAQAICLHPLYVMLPCTLAASLAFMLPVATPPNAIVFSGGLKVSMDMARAGFLNIIIGVL 544
 QY 530 VALVNTNTWGLIIPPEKSPDPWAKEIKNQ 559
 Db 545 AITLSINSWSIPIF-KLDTFPPSWAHSNTSQ 573

RESULT 14
 US-10-718-359-13
 ; Sequence 13, Application US/10718359
 ; Publication No. US20050095240A1
 ; GENERAL INFORMATION:
 ; APPLICANT: MEDICAL COLLEGE OF GEORGIA RESEARCH INSTITUTE, INC.
 ; TITLE OF INVENTION: NaCT AS A TARGET FOR LIFESPAN EXPANSION AND WEIGHT REDUCTION
 ; FILE REFERENCE: 275.0008 0101
 ; CURRENT APPLICATION NUMBER: US/10/718,359
 ; CURRENT FILING DATE: 2003-11-20
 ; PRIOR APPLICATION NUMBER: 60/428,469
 ; PRIOR FILING DATE: 2002-11-22
 ; PRIOR APPLICATION NUMBER: 60/459,441
 ; PRIOR FILING DATE: 2003-04-01
 ; NUMBER OF SEQ ID NOS: 31
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 13
 ; LENGTH: 587
 ; TYPE: PRT
 ; ORGANISM: rat NADCI


```

Query Match      30.0%; Score 868; DB 15; Length 616;
Best Local Similarity 33.5%; Pred. No. 5.7e-72;
Matches 202; Conservative 102; Mismatches 211; Indels 88; Gaps 12;

QY 27 IIPILITLILYGFQDIAEFK---CLWILVIMALLMITETLPIYVYALFPLVFCPLLGL 83
DB 19 VTPELLLPLVI-----LMPKVCSCAYIIMAIYCTEVIPLVATSLMPLLPLFQI 72

QY 84 VNASIVCKQYFTDTIVVFLGGILVALGIEYSNLHTRIALRVIRIVGGSPRRILFVGLMSVS 143
DB 73 LDSRQVCVQYMKDNTNMLFLGGILIVAVAVERNLHKRIARLTLLWVGSAKPARLMLGPMGVT 132

QY 144 TFMGLWISNAGTAMKCPIVKALVNEL-----DTNK--- 174
DB 133 ALLSMWISNTATTAMMPVIVEAILQOMEATSATBAGLEGQGTITNNLNALEDVDTKAVL 192

QY 175 -----IFPVYV-----TQEEPEVBEGBPPHPS-----KIT 199
DB 193 GKKCVAIISTYVKVKEQLINMLTPLKKLEQEQDGLGPIRPQDSAQOEDQERKRLC 252

QY 200 VAFYAGIAVASSIGGLGTILIGTNLVFRGIYTERPPTSTVEITANPMFYISIPLMVIYN 259
DB 253 KAWTLCICVAASIGGTATITGTGPNVLGQMNELFPDSKDIWNFASFAPAFPNMLV-- 310

QY 260 VTLVIIITAPL-ITHMGLFRPNSTKGKIIAEANTNRKLMEDVLRORHIDLGPMSCHEIOMAI 318
DB 311 --MLLEPWLWLFQVYWFSSFKSGWCGGLESKKVEKAALKVLOQEYRKLGLOPLSFAINVL 368

QY 319 AFAPMIVLLITRKPGFVPGWSDLI-----NRKVGSASGLSFTVLVILFALPTQYTFPKYCC 374
DB 369 CFPLLWILWFSRDPGFMPCMLVAVWVEGTEKSVSATVAIVATLLETVPSOKPKNFRS 428

```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2005, 08:29:45 ; Search time 16.7848 Seconds

(without alignments)
3221.603 Million cell updates/sec

Title: US-10-017-479A-3

Perfect score: 2897

Sequence: 1 MAEPGEQRKVLGRCCIFHW.....FPESKSFDPNAKBKINQTKI 562

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 100 summaries

Database : PIR 79:*

1: Piri:*

2: Piri2:*

3: Piri3:*

4: Piri4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	877	30.3	568	JC7911	Na+-coupled citrat
2	830	28.7	593	I46528	sodium/dicarboxyla
3	747	25.8	599	T21613	hypothetical prote
4	725.5	25.0	545	T18694	hypothetical prote
5	687	23.7	595	A47714	Na+/sulfate cotran
6	672.5	23.2	526	S43561	YCR37C homolog K08
7	672.5	23.2	534	G88575	protein K08E5.2 [i
8	642	22.2	539	A88546	protein R107.1 [im
9	617	21.3	520	C99980	hypothetical prote
10	610.5	21.1	539	S30871	hypothetical prote
11	545	18.8	552	F64546	sodium-dependent t
12	536	18.5	548	E71961	probable transport
13	442	15.3	446	F70302	transporter (pho87
14	436	15.1	461	I64080	probable membrane
15	433	14.9	462	B2510	transporter, Nadc
16	425	14.7	471	G81156	transporter, Nadc
17	425	14.7	471	A81948	probable transmemb
18	420.5	14.5	456	F20093	Sodium,sulfate sym
19	396	13.7	487	C92215	probable transport
20	353	12.2	432	C69205	sodium/dicarboxyla
21	337.5	11.6	443	H64383	Na+ transporter -
22	330.5	11.4	478	A82336	probable transport
23	327.5	11.3	867	T40336	probable MSF trans
24	286	9.9	894	S45135	probable membrane
25	283.5	9.8	517	C98940	conserved hypothet
26	258.5	8.9	881	S46633	probable membrane
27	251	8.7	923	MMBY7C	hypothetical prote
28	239	8.2	513	I39534	probable membrane
29	223.5	7.7	450	A10909	hypothetical prote

30	214	7.4	589	2	H84072	sodium/sulfate sym
31	213	7.4	425	2	C75019	transport protein
32	208.5	7.2	610	2	A85870	probable transport
33	208.5	7.2	610	2	H91025	probable transport
34	208.5	7.2	610	2	B65001	probable transport
35	208	7.2	424	2	F71205	hypothetical prote
36	204.5	7.1	608	2	AC0798	hypothetical prote
37	191.5	6.6	428	2	G83803	probable sodium/su
38	189.5	6.5	612	2	S74936	sulfur deprivation
39	183	6.3	618	2	F84409	arsenite transport
40	182.5	6.3	429	2	G70528	probable arsa prot
41	174	6.0	157	2	S36784	mucin - rat (fragm
42	172.5	6.0	428	2	H70528	probable arsa prot
43	170	5.9	449	2	S76839	hypothetical prote
44	165.5	5.7	482	2	F71969	hypothetical prote
45	162.5	5.6	429	2	F87038	probable membrane
46	162	5.6	429	1	B41902	arsenical pump mem
47	161.5	5.6	425	2	F72315	conserved hypothet
48	161.5	5.6	431	2	A11599	arsenic efflux pun
49	161.5	5.6	487	2	B91122	hypothetical prote
50	161.5	5.6	487	2	A85967	hypothetical prote
51	160	5.5	430	2	G89962	hypothetical prote
52	159.5	5.5	610	2	AG0312	probable ion trans
53	158.5	5.5	487	1	Q8ECRS	YgjE protein - Esc
54	158	5.5	461	2	A82185	hypothetical prote
55	156.5	5.4	592	2	D95393	hypothetical prote
56	155	5.4	589	2	F75398	transporter, sodiu
57	154.5	5.3	431	2	AD1237	arsenic efflux pun
58	151.5	5.2	610	2	H83165	probable sodium/su
59	151	5.2	429	1	C41903	arsenical pump mem
60	150	5.2	487	1	B64795	ybdS protein - Esc
61	150	5.2	487	2	G85560	probable membrane
62	150	5.2	487	2	C90710	probable membrane
63	149	5.1	436	2	D98100	conserved hypothet
64	148.5	5.1	436	2	G84024	arsenic efflux pun
65	145	5.0	421	2	G72300	conserved hypothet
66	145	5.0	436	2	C95236	membrane protein [
67	144.5	5.0	435	2	G69779	arsenical pump mem
68	142	4.9	368	2	C71131	hypothetical prote
69	141	4.9	445	2	AD2184	hypothetical prote
70	140.5	4.8	522	2	B64087	Na+/H+-exchanging
71	139.5	4.8	585	2	S69216	sulfur deprivation
72	137.5	4.7	496	2	C69903	proline permease h
73	137.5	4.7	501	2	AE0578	citrate carrier [i
74	137	4.7	512	2	B90050	hypothetical prote
75	135.5	4.7	469	2	T35526	probable integral
76	135	4.7	489	2	AF2970	conserved hypothet
77	135	4.7	489	2	D98312	hypothetical prote
78	134.5	4.6	477	2	G85579	probable membrane
79	134.5	4.6	632	2	F90728	probable dicarboxy
80	134.5	4.6	632	2	A71259	probable dicarboxy
81	133.5	4.6	429	2	AE0406	arsenical pump mem
82	133	4.6	475	2	AI0010	probable membrane
83	132.5	4.6	477	1	B64813	ybhI protein - Esc
84	130.5	4.5	429	1	B25937	arsenical pump mem
85	130.5	4.5	590	2	AG3106	transporter Atu447
86	130.5	4.5	590	2	E98180	transporter, sodiu
87	130	4.5	801	2	A89862	Na+/H+ antiporter
88	128.5	4.4	522	2	A83804	involved in spore
89	126.5	4.4	684	2	B69308	conserved hypothet
90	126	4.3	471	2	F71543	probable dicarboxy
91	125	4.3	484	2	AE0887	probable membrane
92	124.5	4.3	417	2	AF2992	arsenical pump mem
93	124.5	4.3	417	2	D98291	probable arsenical
94	124.5	4.3	388	2	AC3587	sulfur deprivation
95	124	4.3	541	2	AF1541	antibiotic resista
96	123.5	4.3	440	2	G84069	hypothetical prote
97	123.5	4.3	473	2	B95376	probable Arc02 arg
98	123.5	4.3	700	2	H82428	iron(III) ABC tran
99	123	4.2	369	2	A75166	hypothetical prote
100	123	4.2	469	2	C90446	permease, multidru

ALIGNMENTS

RESULT 1

JC7911
Na+-coupled citrate transporter NaCT - human
C:Species: Homo sapiens (man)
C>Date: 31-Mar-2003 #sequence_revision 31-Mar-2003 #text_change 14-Jul-2003
C/Accession: JC7911
R;Inoue, K.; Zhuang, L.; Ganapathy, V.
Biochem. Biophys. Res. Commun. 299, 465-471, 2002
A:Title: Human Na+-coupled citrate transporter: Primary structure, genomic organization, or the synthesis of fatty acid and chol 17p12-13
A:Reference number: JC7911; MUID:22334959; PMID:12445824
A:Accession: JC7911
A:Molecule type: mRNA
A:Residues: 1-568 <INO>
A:Cross-references: GB:AY151833
C:Comment: This transporter classified as a tricarboxylate transporter represents the first the cellular entry of citrate by a process energized by the electrochemical Na+ gradient
C:Genetics:
A:Gene: nact

Query Match 30.3%; Score 877; DB 2; Length 568;
Best Local Similarity 35.7%; Pred. No. 2.1e-59;
Matches 198; Conservative 99; Mismatches 218; Indels 40; Gaps 10;
QY 27 IIPILITPILYIGFQDMARF-KCLMLIVTMALLWITETPIYVTALFPLVFCPLGLLVN 85
DB 19 VTPLLPLPLV-----LMPAKFVRCAVYIILMAIYVCTEVIPLAVTSLMPVLLPFLQILD 74
QY 86 ASIVCKOYFTDTIVVFLGGILVALGIEYSNLHTRIALRVIRVIGGSPRRLFVGLMSVSTF 145
DB 75 SRQVCVQYKMDTNLFGGLIVAVAVERNLHKRIARLTLWGAQPARMLGMFGVTAL 134
QY 146 MGLWISNAGTAMMCPIVKALVNELDTN-----KIPPVYMTQEEPV-E 188
DB 135 LSMWISNTAITAMVPIVAILQMEATSAATEAGLELVDKRAKELPGSQVIFEGPILG 194
QY 189 EGEPPHPSKITVAYAGIAYASSIGGLTGLTGNLVFRGIYTERPPTSTVITFANFM 248
DB 195 QOEDQERKRCLCKAMTLCICYAASIGGTATLTGTGPNVVLVGMNELFPDSKDLVNASWF 254
QY 249 FYSIPLMIVNVTVIIAFLITMGLFRPN-SKTGKIIAEANTNRKLMEDVLRORHIDL 307
DB 255 AFAPFNMLV-----MLFANLWLFQVYRNFNFKSWGCGLESKKNKAAKVLQEEYRKL 310
QY 308 PMSCHEIQMAIFAFAMIVLLITRKPGFVPGWSDLI-----NRKVVGSSAGSLSFVILLIFAL 363
DB 311 PLSFAEINVLICFLLVILFSDRDPGMPGWLTVAWVEGETKYVSDATVAIFVATLLFIV 370
QY 364 PTQTFYKYCC-----GKGPPTAQAIADAILSWEVYVLRNIPKGLLFLGCGGFALAVASRET 418
DB 371 PSQPKFNFRSQTBEERKTPFP-----PPLLDKWTQEKVPMGIVILLGGGFALAKGSEAS 427
QY 419 GLNIMISKAMOVILGLPNVQVSTFVLNPFSAFNANVAVANVILPILCMSLALHLP 478
DB 428 GLSVMGKQMEPLHVPAAITLILSVLVFTECTSNVATTLFLPFPASMSRSIGLNP 487
QY 479 LILTLPACLGISVMYFPLPVSTPPNNAIVTQYAHIKTKYFACCGIVPTTIIGISVALVNTW 538
DB 488 LYIMLPCTLSASFAPMLPVATPPNNAIVFTYGHKLVDAMVKVTGVTMNIIGVFCVFLAVNTW 547
QY 539 GLIIFPESKSPDWA 553
DB 548 GRAIF-DLDHFPDWA 561

RESULT 2

146528
sodium/dicarboxylate cotransporter - rabbit (fragment)
C:Species: Oryctolagus cuniculus (domestic rabbit)
C>Date: 14-Feb-1997 #sequence_revision 14-Feb-1997 #text_change 09-Jul-2004

C/Accession: 146528

R;Pajor, A.M.
J. Biol. Chem. 270, 5779-5785, 1995
A:Title: Sequence and functional characterization of a renal sodium/dicarboxylate cotrans
A:Reference number: 146528; MUID:95197598; PMID:7890707
A:Accession: 146528
A>Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-593 <PAJ>
A:Cross-references: UNIPROT:Q28615; EMBL:U12186; NID:g758383; PIDN:AAA99666.1; PID:g75838
C:Superfamily: sodium/sulfate cotransporter

Query Match 28.7%; Score 830; DB 2; Length 593;
Best Local Similarity 34.8%; Pred. No. 8.7e-56;
Matches 200; Conservative 107; Mismatches 199; Indels 68; Gaps 14;
QY 28 IPIITLPIIYGFQDMARF-KCLMLIVTMALLWITETPIYVTALFPLVFCPLGLLVN 87
DB 22 ISLLPLPLV-----PRKEAYCAYAIILMALFWCTDALPLAVTALLPLCLFPMGIMEAS 76
QY 88 IYCKOYFTDTIVVFLGGILVALGIEYSNLHTRIALRVIRVIGGSPRRLFVGLMSVSTF 147
DB 77 EYGLEVLKDTNVLFPGGLLATAVEHNLHKRIARLRLVLLTGVRRALLILGFMVTAFLS 136
QY 148 LWISNAGTAMMCPIVKALVNELDTNKI-----FPVYMTQEEB----- 185
DB 137 MWISNTASTAMVPIAHAVLQELNNTQSNVERGSDNPTPELOEPSQKETSVDKXNGQ 196
QY 186 -----PVEEGEPHPSKITVAFYAG-----TAYASSIGGLTGLTGNLVFRGIYTER 234
DB 197 AQPLPAVPLESGE--HMQEQRLRFSGMSLSCYCSASIGIATLTGTTNVLVLOGQNTSL 254
QY 235 FPTSVEITFANFMFYISIPLMIVNVTVIIAFL---ITHMGL-FRPNSTKTKIIAEANT 290
DB 255 FQPNPNVNFASFGEFAPFIMVI-----LILLSMLQILFLGINFRKN--FGIREQEHEQ 308
QY 291 NRK-LMEDVLRORHIDLGPMSCHEIQMAIFAFAMIVLLITRKPGFVPGWSDLI-----NR 344
DB 309 QRKQAYRVIOQYRLGLGPMSEFAEKAVFLFVLVLLWFTREPGFPHGWNLVFSDASGR 368
QY 345 KVVGSAGSLSFVILLIFALPTQYTFPKYCCG-----KGPPTAQAIADAILSWEVYVLRNIPW 399
DB 369 VMVDSGASILGVFLFWFSKIP-----GLTQDPDNFGRLLKAPALLNKLNVKMPW 422
QY 400 GLLFLGCGGFALAVASRETGLNIMISKAMQVILGLPNVIVQVSTFVLNPFSAFNANVV 459
DB 423 NVILLGGVYALAKGSEESGLSQWLGNLMLPLQHVPPATVFIICLLVATFTECTSNAAT 482
QY 460 ANIVLPILCMSLALHPLILTLPACLGISVMYFPLPVSTPPNNAIVTQYAHIKTKYFACC 519
DB 483 TTLLPLILASMAQAICLHPLYVMLPCTLASSLAFMLPVATPPNNAIVFSGGLRVSDMARA 542
QY 520 GIVPTTIIGISVALVNTWNTWGLIIFPESKSPDWA 553
DB 543 GIMNIIGVIMLAINSGVPMF-QLHTFPSWA 575

RESULT 3

T21613
hypothetical protein F31P6.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T21613
R;Percy, C.
submitted to the EMBL Data Library, March 1996
A:Reference number: Z19449
A:Accession: T21613
A>Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-599 <WIL>
A:Cross-references: UNIPROT:Q93655; EMBL:Z69884; PIDN:CAA93752.1; GSPDB:GN000028; CESP:F3J
A:Experimental source: clone F31P6
C:Genetics:

A:Gene: CESP.F31F6.6
A:Map position: X
A:Introns: 39/3; 142/2; 207/2; 255/2; 285/2; 397/1; 437/3; 493/2; 537/3
C:Superfamily: sodium/sulfate cotransporter

Query Match 25.8%; Score 747; DB 2; Length 599;
Best Local Similarity 30.1%; Pred. No. 2e-49;
Matches 174; Conservative 114; Mismatches 223; Indels 68; Gaps 10;

Qy 23 KASIII---PLITLIIYGFQDMAEFKCLWLIIVTMALLMITETLPIYVTFALPFLVFCP 79
Db 15 KQSFVIGALLIFSLLMFVGDHGLQAKCLYCVAVMGVYVFEALPLAIFAPIMILFP 74
Qy 80 LGLVNASIVCKQYPTDITIVPLGLGIVAGLIEYSLNHLRIALRVIRVGGSPRLFLVGL 139
Db 75 LFGIMRSEVAYLPDCTFLFMGLMVALAVEKELHARVALFLVLTGVGSEPARVMAGF 134
Qy 140 MSVSTFMGLWISNSAGTAMWCPIVKALVNELDN-----KIPFVYMTQ 182
Db 135 MGVTFLSMISNTATTALMPILQSVITELVSNHRMEDVALCEAHNSRKHSVGMRR 194
Qy 183 EEPVEEGE-----PHPSKITVAFVAGIAYASSIGLGLTIGTNLVRGIYT 232
Db 195 LSLPNNENIKREENDTAMSPREQMAKGLMLSVCFSANIGGAATITGASNLVLVGQLN 254
Qy 233 ERFPSTVEITFANPMFYSIPLMIVNVTVIIAFLITHMGLFRPNSTKGIIEAENTNR 292
Db 255 ELFPGADTGVNFLSWLIEAFPMVFCCLIVCWCVLV-----LYLRDAPKGSII---VTR 305
Qy 293 KLMDVLQRHIDLGPMSCHIEIOMIAIAFAMIVLLITRKPGFVQWSDL----- 341
Db 306 K-----LOQKNELHAFSAEMAVIFCFALLLVILREPOVPGWGMFKDELVFKSLT 360
Qy 342 -----INRKVGSASGLSFVLLIFALPTQYTFKCYCGKGP- -TAQIDAISLWVYL 394
Db 361 EKQNTLHTRFVSDATSAFVILLTFLPE-----KLPSSRGSSBQRKASSGLDWTQV 415
Qy 395 RNIWMLGLFLGGGFALAVASRETGLNIMISKAMQVLGLPNVIVQSTFVLANFFSAFN 454
Db 416 DRFPWSVFLGGGFALAAVKGESGLSHDIGAIMRYLDVFNHNIIMLCIIISVLTNVC 475
Qy 455 ANVVANIVLPILCMSIALELHPILITPLACLGISMYVFLPVSTPPNAIVTQYAHIKTK 514
Db 476 SNTVIASIFIPVABLARSLEIDPLNFMPLPTISASFAFLPLPVATPPNAIVFSSGYLKV 535
Qy 515 YFACCGIVPTIIGISVALVNTWGLIIFPSKSPDWA 553
Db 536 DMFVSGLCVTILGCVVLSMLNMLLWAGFVF-NLHLFPQWA 573

RESULT 4
Tl8694
hypochemical protein B0285.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C:Accession: Tl8694
R: Sulston, J.
submitted to the EMBL Data Library, June 1994
A:Reference number: Z19007
A:Accession: Tl8694
A:Status: preliminary;
A: Molecule type: DNA
A:Residues: 1-545 <WIL>
A:Cross-references: EMBL:Z34533; PIDN:CAA84299.1; GSPDB:GN000021; CESP:B0285.6
C:Genetics:
A:Gene: CESP:B0285.6
A:Map position: 3
A:Introns: 6/3; 56/3; 95/2; 107/2; 160/2; 175/1; 211/2; 287/2; 353/1; 393/3; 494/3
C:Superfamily: sodium/sulfate cotransporter

Query Match 25.0%; Score 725.5; DB 2; Length 545;
Best Local Similarity 31.5%; Pred. No. 7.9e-48;

Matches 174; Conservative 114; Mismatches 212; Indels 53; Gaps 13;
Qy 43 DMAEFKCLWLIIVTMALLMITETLPIYVTFALPFLVFCPLGLVNASIVCKQYFTIVVFL 102
Db 4 ESTEGKCAFVILITWSCYVVAEVLAVTSFIPMIALPFLGIVSIVKAVPKYFADTNIVFF 63
Qy 103 GGLIVAGLIEYSLNHLRIALRVIRVGGSPRLFLVGLMSVSTFMGLWISNSAGTAMWCPI 162
Db 64 NSLMSLAVBECQLHKRIALKMLTYVGTTRPHLMAGFMIIITPSISLWISDACCALMAPI 123
Qy 163 VKALV-----NELDNKIFPVYMTQEEBPVE-EGEPHPHSKITV----- 200
Db 124 AYALLEIIMPQPEKEKEIEVWKIF-----DKEDPEEKEKKLDTLSRLSVDRGRICK 178
Qy 201 AFYAGIAYASSIGLGLTIGTNLVRGIYTERFTSTVEITFANPMFYSIPLMIVNV 260
Db 179 CMLLVAHASLIGGTGTINSTGNLIIFRDNIEKNFPNEDHGISYLSWMAFAPPMIYFMP 238
Qy 261 T---LVIIALITH--MGLFRPNSTKGIIEAENTNRKLMEDVLQRHI-----DLGMS 310
Db 239 SSWFIVQLQFLGPRHLMGMPREPTETEK-----QEEVAKRAVWKSYDQLGEMT 287
Qy 311 CHEIQMAIAFAFMIVLLITRKPGFVQWSDLINR-KVVGSAAGLSFVLLIFALPTQYTF 369
Db 288 WAEXSTVIFVLAVLSWSSDPKVPWSDLFRGYVTDSCGL-VAVLFLFIPWKKPD 346
Qy 370 FKYCCGKGPFTAQIDAISLWVYVLRNIPWGLLGLGGFALAVASRETGLNIMISKAMQ 429
Db 347 FRFRKDKSRSPVQEPDLIDWCVRFRFPWSIILLGAGFAISDAVRVSGLSLIACSLN 406
Qy 430 VLIG-LPNIIVQSTFVLANFFSAFNANVVANIVLPILCMSIALELHPILITPLACLG 488
Db 407 STISKMPFFVWQIILSIWVVMTEFSTNSATASIFIPISFKMAEAVGAHPLFSIPTAIG 466
Qy 489 ISMYVFLPVSTPPNAIVTQYAHIKTKYFACCGIVPTIIGISVALVNTWGLIIFPSK 548
Db 467 PSFSLMPLMATPANAIYETKTIRMDVSCGVFLNFCIAITAINMTWAFWLF-NMGT 525
Qy 549 FPDWA-KEIKNOT 560
Db 526 YPDVALRHATNMT 538

RESULT 5
A47714
Na+/sulfate cotransporter, renal - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 02-Jun-1994 #sequence_revision 02-Jun-1994 #text_change 09-Jul-2004
C:Accession: A47714
R: Markovich, D.; Forgo, J.; Stange, G.; Biber, J.; Murer, H.
Proc. Natl. Acad. Sci. U.S.A. 90, 8073-8077, 1993
A:Title: Expression cloning of rat renal Na+/SO4(2-) cotransport.
A:Reference number: A47714; MUID:93376745; PMID:7690140
A:Accession: A47714
A:Status: preliminary
A: Molecule type: mRNA
A:Residues: 1-595 <WAP>
A:Cross-references: UNIPROT:Q07782; GB:L19102; NID:g310182; PIDN:AAA41677.1; PID:g310183
C:Superfamily: sodium/sulfate cotransporter

Query Match 23.7%; Score 687; DB 2; Length 595;
Best Local Similarity 28.9%; Pred. No. 7.7e-45;
Matches 173; Conservative 116; Mismatches 220; Indels 90; Gaps 14;

Qy 27 IIPILITPLIYGFQDMAEFKCLWLIIVTMALLMITETLPIYVTFALPFLVFCPLGLVNA 86
Db 20 VLVLPLPLIIRS-----KEACAYILFVIAETFEALPLSITALLPGLMFPFMGIMSS 74
Qy 87 SIYCKQYFTIVVPLGLGIVAGLIEYSLNHLRIALRVIRVGGSPRLFLVGLMSVSTEM 146
Db 75 THVASAYFKDFHLLIGVICLATSIKKNLHKLRIALRMVMVMVPAWLTIGFMSSTAF 134
Qy 147 GLWISNSAGTAMWCPIVKALVNELDN-----KIPFVYMTQEEBP 186

```
Db 135 SMWLSNTSTAAMVMPIVEAAVAQTSAAEABAEATQMTYFNESAQAQGLEVDETIIQGETNE 194
QY 187 VEGEPP-----HPSKITVAFVAG-----TAYASSTGGGLG 216
Db 195 RKEKTKPALGSSNDKGKVSXKMETEKNTVTGAKYRSKKOHMMCKMLCLCAYSSSTIGGLT 254
QY 217 TLIGTGNLVRFGIYTERFPTSTVEITFANMFYSIPLMIVNVTLVIIAFLITH---MG 273
Db 255 TIIGTGNLVRFGIYTERFPTSTVEITFANMFYSIPLMIVNVTLVIIAFLITH---MG 273
QY 274 L-FRPNKTKGKIAEANTNRKLMEDVLRQRHIDLGPMSCHEIOAIAFAPMIVLLITRK 332
Db 310 FNFKEPKCKG--TKTLKEKACAEVTKQBYEKLGPMPYQOIVTLVIFIVMALLWFSRDP 366
QY 333 GFVPGWSDL-----INRKVGSASGLSFIVLLIFALP-----TQYTFKYCCGKGPFTA 381
Db 367 GFVTGWSVLFSEYPGYVTDSTVALVAG-----ILFFLIPAKKLTKMT-----STGDI 415
QY 382 QAIDAILSWEVLRNIPWGLLFLGGGFALAVASRETGLNIMISKAMQVLIPLNIVQ 441
Db 416 FDYSPLTWKEFQSPFMDTALVGGGFALADGQVSGLSWIGSKLSPGLSLPVMIL 475
QY 442 ITFVLNFFSAFNANVVANIVLPILCEMSLALELHPLILTLPAACIGISMVYFLPVSTPP 501
Db 476 ISSLIVTSLEVASNPATITILPILSPLAEATHVNPILHLLPSTLCTSPAFLLPVANPP 535
QY 502 NAIYVQAHITKTKYFACCGIVPTIIGISVALVNTNWTGLIIFPESKFPDPAKEIKNOT 560
Db 536 NAIYVSYGHLKVIDMVRKAGLVNIGLVAVVLMGMFTWIEPMP-NLHYPSPWAPDVNQT 593

RESULT 6
S43561
YCR37C homolog K08E5.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 07-Sep-1994 #sequence_revision 10-Nov-1995 #text_change 24-Nov-1999
C;Accession: S43561
R;Kershaw, J.
submitted to the EMBL Data Library, March 1994
A;Reference number: S43561
A;Accession: S43561
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-526 <KER>
A;Cross-references: EMBL:Z30974
C;Genetics:
A;Introns: 33/3; 122/2; 169/2; 214/2; 253/2; 481/3
C;Superfamily: sodium/sulfate cotransporter
```

```
Query Match 23.2%; Score 672.5; DB 2; Length 536;
Best Local Similarity 29.1%; Pred. No. 8.6e-44;
Matches 154; Conservative 116; Mismatches 206; Indels 53; Gaps 9;

QY 46 EFKCLMWLVMTALLWITETLPIYVTFALFPLVFCPLGLVNASIVCKQYFTDTIIVFLGGL 105
Db 34 EWKCAVCVCIIVYMWSEVMPPLAVTAMLPVLPVLGVLDANTTAKEYNDTNFLFIGL 93
QY 106 IVALGIEYNLHTRIALRVIRIVGSPRRPLFVGLMSVSTFMGLWISNSAGTAMMCPIVKA 165
Db 94 IMAAAVEKCDLHERVALSVLRVCGSEPKWMLGFMVTVALLSSFSISNTATTAMVPVIGS 153
QY 166 LVNELDTNKLFPVYMTQEEPEVEGEPHPS-----KITVAFYAGIAYASSIGGL 215
Db 154 VVOQLISS-----FOHPTNGERGLGCKKMATGLVLISICFAANIGGT 196
QY 216 GTLIGTGNLVRFGIYTERFPTSTVEITFANMFYSIPLMIVNVTLVIIAFLITHMGLF 275
Db 197 GTATGTPSNLVMQLGQSLALFPKVDGSLNVTWIFFAYPLML-----LCLFVAMWTLVSFF 251
QY 276 RPNKTKGKIAEANTNRKLMEDVLRQRHIDLGPMSCHEIOAIAFAPMIVLLITRKPGFV 335
Db 252 -----LRDAPEKDEAVTEMLKTRYNELPRMTYAEKSVFVFCILLSLWVFRNPGVV 302
```

```
QY 336 PGWSDLINRKVGSASGLSFIVLLIFALPQTQ-----YTFPKYCCGKGPFTAQAIDAILSWE 391
Db 303 PGGFVFFKGGAYDATSAMIVAFLLFVLPSEPDLATYIK----KEDLKKRG--CLMDWK 356
QY 392 YVLRNIPWGLLFLGGGFALAVASRETGLNIMISKAMQVLIPLNIVVQSITFVLNFFS 451
Db 357 TMOETFPWVLLGGGFALAAAGVKESGLSLIGNSLSIEHLPLWILQQLTMLIAMVIT 416
QY 452 AFNANVVANIVLPILCEMSLALELHPLILTLPAACIGISMVYFLPVSTPPNAIVTQYAH 511
Db 417 NICSNTVTASIFVIVATLAQAGHHPTMLPTTLASSFAFIPFVGTTPNALVFGSGMW 476
QY 512 KTKYFACCGIVPTIIGISVALVNTNWTGLIIFPESKFPDPAKEIKNOT 560
Db 477 KYSDMAFVGIIISLELLVLTVMNSIAYTLPL-LLEFFPTWA-IIANST 523

RESULT 7
G88575
protein K08E5.2 [imported] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-May-2001
C;Accession: G88575
R;anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A;Reference number: A75000; MUID:99069613; PMID:9851916
A;Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/projects/C_eleg
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A;Accession: G88575
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-534 <STO>
A;Cross-references: GB:chr_III; PIDN:CAA83225.1; PID:G3878357; GSPDB:GN00021; CBSP:K08E5.
C;Genetics:
A;Gene: K08E5.2
A;Map position: 3
C;Superfamily: sodium/sulfate cotransporter
```

```
Query Match 23.2%; Score 672.5; DB 2; Length 534;
Best Local Similarity 29.1%; Pred. No. 8.6e-44;
Matches 154; Conservative 116; Mismatches 206; Indels 53; Gaps 9;

QY 46 EFKCLMWLVMTALLWITETLPIYVTFALFPLVFCPLGLVNASIVCKQYFTDTIIVFLGGL 105
Db 42 EWKCAVCVCIIVYMWSEVMPPLAVTAMLPVLPVLGVLDANTTAKEYNDTNFLFIGL 101
QY 106 IVALGIEYNLHTRIALRVIRIVGSPRRPLFVGLMSVSTFMGLWISNSAGTAMMCPIVKA 165
Db 102 IMAAAVEKCDLHERVALSVLRVCGSEPKWMLGFMVTVALLSSFSISNTATTAMVPVIGS 161
QY 166 LVNELDTNKLFPVYMTQEEPEVEGEPHPS-----KITVAFYAGIAYASSIGGL 215
Db 162 VVOQLISS-----FOHPTNGERGLGCKKMATGLVLISICFAANIGGT 204
QY 216 GTLIGTGNLVRFGIYTERFPTSTVEITFANMFYSIPLMIVNVTLVIIAFLITHMGLF 275
Db 205 GTATGTPSNLVMQLGQSLALFPKVDGSLNVTWIFFAYPLML-----LCLFVAMWTLVSFF 259
QY 276 RPNKTKGKIAEANTNRKLMEDVLRQRHIDLGPMSCHEIOAIAFAPMIVLLITRKPGFV 335
Db 260 -----LRDAPEKDEAVTEMLKTRYNELPRMTYAEKSVFVFCILLSLWVFRNPGVV 310
QY 336 PGWSDLINRKVGSASGLSFIVLLIFALPQTQ-----YTFPKYCCGKGPFTAQAIDAILSWE 391
Db 311 PGGFVFFKGGAYDATSAMIVAFLLFVLPSEPDLATYIK----KEDLKKRG--CLMDWK 364
QY 392 YVLRNIPWGLLFLGGGFALAVASRETGLNIMISKAMQVLIPLNIVVQSITFVLNFFS 451
Db 365 TMOETFPWVLLGGGFALAAAGVKESGLSLIGNSLSIEHLPLWILQQLTMLIAMVIT 424
QY 452 AFNANVVANIVLPILCEMSLALELHPLILTLPAACIGISMVYFLPVSTPPNAIVTQYAH 511
```

Db 425 NICSNTVTASIEVPVIVATLAQAGHPTLMPTTLASSFAPFPVGPVPPNAIVFGSGMV 484
Qy 512 KTKYPACCGIVPTIIGISVALVNTWGLIIPPEKSPDPDAKEIKNOT 560
Db 485 KVSDFAFVGGIISLELLVLTVMYNSIAYLTLP-LLEPTWA-IIANST 531
RESULT 8
A88546
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
R:Accession: A88546
R:Anonymous, The C. elegans Sequencing Consortium.
C:Title: Genome sequence of the nematode C. elegans: a platform for investigating biological processes
A:Reference number: A75000; PMID:9906913; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_elegans/
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and Science 282, 2012-2018, 1998
A:Accession: A88546
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-539 <STO>
A:Cross-references: UNIPROT:P32739; GB:chr_III; PIDN:CAA78468.1; PID:G3879033; GSPDB:GNO
A:Note: predicted using GeneFinder
C:Genetics:
A:Gene: R107.1
A:Map position: 3
C:Superfamily: sodium/sulfate cotransporter
Query Match 22.2%; Score 642; DB 2; Length 539;
Best Local Similarity 29.6%; Pred. No. 1.9e-41;
Matches 175; Conservative 109; Mismatches 209; Indels 98; Gaps 20;
Qy 3 EPGEORKEVLCRCIFHWGKASII--PLITPLIYIGFQTDMAEFKCLMLIVTMALLW 60
Db 2 KPSPQRTLI-----KCLLVLLGLVAVPLPFFG-----PEYRCLFSIIFLSTY 45
Qy 61 ITETPLIVVTALFPLVFCPLGLVNASIVCKQYFTDTIIVFLGGLIYALGIEYNSLHTRI 120
Db 46 IGEAFPGVTSFLPALLYPILOIYVPSKQISPVYFQDSIVLFWCTLMMAVAEATGLHRI 105
Qy 121 ALRVIRIVGSGPRRLVFLGMSVSTFMGLWISNSAGTAMMCPVVKALVNELTNKFPPVM 180
Db 106 ALKLLTKVGAK-----QPVSEF-----VSDTACTALMCTPAVALLMSD-----AV 147
Qy 181 TOEEPEVEEGPPHPSKITV-----AFVAGIAYASSIGGLGTLTGTG 222
Db 148 QHLKEDHRKPKPP-PDDATVAEKLRIIDMTPODAGFCALILACAHASLIGGTAIITSTG 206
Qy 223 TNLVFRGIYTEREPTSTVEITFANFMFYSIPLMIVNVT--LVIIAFLI---THMGLPRP 277
Db 207 PNLVFRNIHKRYPEGQVTMYLQMWFAIDPMPFYLLASVIIIVCYPMGSGTFAFMFER 266
Qy 278 NSKTGKIIEANTNRKLMEDVLRQRIIDLGPMSCHIEQMAIAFAFMIVLLITRKPGFVPG 337
Db 267 PSK-----EEAHL-KKLIENKIQTMYEDLDGVSNGEKSVFVFILLIGSWISRDPGTTP 320
Qy 338 WSDLINRK--VVGASGLSFVLLIIPALPTQYTFKVCYCCGKGPPTAQAIID---AIIISWEY 392
Db 321 WGDLLPHRNFTSDSVSGV-LISCILFVMP-----KDPE--DPIDPMAPILKWT 366
Qy 393 VLRIIPWGLLELGGGFALAVASRETGLNTIMSKAMQ-VLIGLPIVIVQSIFFVLNFFS 451
Db 367 MKSKFESWCTLLIGAGVAISEGVDKSLRSLICGMKNIFVGMSSLPQLTFTVTIIVMT 426
Qy 452 AFANANVVANIVLPILCEMSLALHLLPLTLPACLGISVMYFLPVSTPPNAIVTQVAHI 511
Db 427 EFASNVSTGSIFFIISLGVAESMGVHPLYLALPTTVACSFAPMLPISTPPNAVVDYTKVI 486
Qy 512 KTKYFACCGIVPTIIGISVALVNTW-----GLIIFP-----ESKSPF 550

Db 487 SMVEMIVCGFLNLITSLANMTWTYFIFSLNIPFENIVIVISSENSYP 537
RESULT 9
C89980
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
R:Accession: C89980
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguchi, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiratsugu, K.
A:Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.
A:Reference number: A89758; PMID:21311952; PMID:11418146
A:Accession: C89980
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-520 <KUR>
A:Cross-references: UNIPROT:Q99SX1; GB:BA000018; PID:G13701709; PIDN:BAB43002.1; GSPDB:G
A:Experimental source: strain N315
C:Genetics:
A:Gene: SA1732
C:Superfamily: sodium/sulfate cotransporter
Query Match 21.3%; Score 617; DB 2; Length 520;
Best Local Similarity 29.4%; Pred. No. 1.5e-39;
Matches 155; Conservative 102; Mismatches 197; Indels 74; Gaps 13;
Qy 26 IIPILITPLIYIGFQTDMAEFKCLMLIVTM--ALLWITETPLIYVVTALFPLVFCPLGL 83
Db 37 ILGPELLLTLLFPHPQDLPWKGVYVLAITLWIAITWITEAIPAAITSLPVLPLGLH 96
Qy 84 VNASIVCKQYFTDTIIVFLGGLIYALGIEYNSLHTRIIVIRIVGSGPRRLVFLGMSVS 143
Db 97 LTPEQVSEYGNIIIFLGLGFILAIAMRNWNLHTRVALTIINIGASTSKILLGFMVAT 156
Qy 144 TFMGLWISNSAGTAMMCPVVKALVNELTNKFPPVMYTOEEPEVEEGPPHPS--KITVA 201
Db 157 GFLSMFVSNATAVMIMIPIGLAI-----KEAHDLOEANTNQTSTQKFEKS 202
Qy 202 FYAGIAYASSIGGLGTLTGTNLFVFRGTYTERPPTSTVEITFANFMFYSIP-LMVIIVN 260
Db 203 LVLAIGVAGTIGGLTIGTLPPLIILKGQYMQHFGH---EISFAKMMIVGIPTVIVLLGI 259
Qy 261 TLVIAFL-ITHMGLFEPNSKTGKIAEANTNRKLMEDVLRQRIIDLGPMSCHIEQMAIA 319
Db 260 TWLYLRVAVFRHDLUKYLPQGT-----LIQKLDLGLKMKYBEKVQTI 303
Qy 320 FAFMIVLLITRKPGFVPGWSDLINRKVVGSGLSFIVLLIFALPTQVTFPKYCCGKGP 379
Db 304 FVLASLLWITRE-FLKKWE--VTSSVADGTIAI-FISILLFIIPAKNT----- 348
Qy 380 TQAIDAILSWEVYLRNIPWGLLFLGGGFALAVASRETGLNTIMSKAMQVILGLPIVIV 439
Db 349 --EKHRRIIDWE-VAKELPGLVILFGGLALAKGISEGLAKWLGQLKSLNGVSPILI 405
Qy 440 QSIFFVLNFFSAFNANVVANIVLPILCEMSLALHLLPLTLPACLGISVMYFLPVST 499
Db 406 VITITFVLEFTEVTSNTATATMLPILATLSVAVGVHPLLLMAPAANAANCAYMLPVGT 465
Qy 500 PPNAIVTQVAHIKTKYPACCG-----IVPTIIGISV 530
Db 466 PPNAIFGSGKISIKQWASVGFVWNLSAIIILVVYVMPVIVLGIDI 513
RESULT 10
S30871
C:Species: Caenorhabditis elegans
C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 21-Jan-2000
R:Thomas, K.

Db 68 VLMGIWMTAEADLPATALLPLVFNVSVDQFSSVSASYPPIFLFMGGFIALSWMOK 127
Qy 114 SNLHTRIALRVIRIVGSSPRRLFVGLMSVSTFMGLMISNAGTAMMCPT-----VKAL 166
Db 128 WNLHTRIALSILLVGTSPRRLFGPIATGFLSMWNSNTATVMFPVGMVQLQVAKL 187
Qy 167 VNELDTNKIPVYMQE-----EPPVEGEPPHPSKITVAFYAGIA 207
Db 188 VGENASNAF--YQEEITKAHGIMSNIHVHKDKIAQVIOEKTIYRTNFSICLMLGIA 245
Qy 208 YASSIGGLTGLTGTNLVFRGIYTERPPTSTVEITPANFMYSPILMVIWVTLVIA 267
Db 246 YSASIGSLGTLTGPNNALLAGYMKTF--NIEIDFAQMVFPGTFLAPI-----MLILAW 298
Qy 268 LITHMGLFRPNSKT--GKIIAEANTNRKLMEDVLRQRHIDLGPMSCHEIQMAIAFAPMI 324
Db 299 LLITVYIFPLKIEIPGKGVKSEL--KLRR--LSQAEISGVIF--ILASLGWIFLD 352
Qy 325 VLLITRKPGFVPGWSDLINRNVGSGASLSFTVLILFALPTQYTFPKYCCGKGPFTAQAI 384
Db 353 TIL-----KSWGIKIDKIDSVIANGVS--ALLFILPANH-----QG----- 386
Qy 385 DAILSWEVLRNIPWGLLFLGGSFALAVASRETGLNIMISKAMQVLGLNIVVQSITF 444
Db 387 DRLLDWG--VAKKLPWDVLLFLGGLSALSQAQSKTGLSLWIGHLVSGFHLPLFIIVMVT 445
Qy 445 VLANFSAFNANVAVANIVLPIL--CEMSLALELH--PLILTPACLGISMVYFLPVSTPP 501
Db 446 LMVIELTETSNTATAAFPLVIGVGAMGVESHQSLLITIPVALSATCAFMPLVATPP 505
Qy 502 NAIVTQYAHIKTKYFACCGIVPTIIGISVALVNTWGLI 541
Db 506 NAIAYGSGYVKITDMIKAGLWNLVG--VVLISAFSYCLV 543
RESULT 13
F70302
transporter (Pho87 family) - Aquifex aeolicus
C:Species: Aquifex aeolicus
C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004
C:Accession: F70302
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ovi
V.
Nature 392, 353-358, 1998
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A:Reference number: A70300; PMID:98196666; PMID:9537320
A:Accession: F70302
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-446 <AQF>
A:Cross-references: UNIPROT:O66449; GB:AE000670; NID:g2982779; PIDN:AAC06409.1; PID:g298
A:Experimental source: strain VFS
C:Genetics:
A:Gene: trns
C:Superfamily: probable transporter MJ0672
Query Match 15.3%; Score 442; DB 2; Length 446;
Best Local Similarity 26.5%; Pred. No. 3.1e-26;
Matches 138; Conservative 85; Mismatches 179; Indels 118; Gaps 14;
Qy 30 LITPLIYGFQDMAEPKCLMWLIVTMALLMITETPLIYVVTALFPLVFCPLLGLVNASIV 89
Db 18 LITSP-----FEGNVN--KGIAILLAAAILMITEALPLVPTALLIPVSGVLLGVFDVKT 70
Qy 90 CKQYPTDITVIVPLGLLIVAGLIEYNSLHTRIALRVIRIVGSSPRRLFVGLMSVTFMGLW 149
Db 71 LSYFAHPLIFLFFGGLVLAVALSKYQIDEYIAHKIVSVAQKFLPSVFLMLATSLISMW 130
Qy 150 ISNSAGTAMMCPI---VKALVNELDTNKIPVYMQEPEEPVEGEPPHPSKITVAFYAGI 206
Db 131 ISNTSTTAMPLALGILAGVRETEREKVFPVL-----LGI 167

Qy 207 AYASSIGGLTGLTGTNLVFRGIYTERPPTSTVEITPANFMYSPILMVIWVTLVIA 266
Db 168 YSASVGGIGTLVSGPPNGIAAGI-----LGLSPDWLKFQGVFLFLFLLFAIL 218
Qy 267 FLITHMGLFRPNSKTGKIIAEANTNRKLMEDVLRQRHIDLGPMSCHEIQMAIAFAPMVL 326
Db 219 FLV-----FRPTS-----DLKVERVOEIKFETPORVILV 248
Qy 327 LI-----TRK--PGF--VPGWSDLINRNVGSGASLSFTVLILFALPTQYTFPKY 373
Db 249 LIFLTALAWIFSCKIAPIFEVKYFDIV-----VALLAVVALFIRL----- 291
Qy 374 CGKGPFTAQAIADAILSWEVLRNIPWGLLFLGGSFALAVASRETGLNIMISKAM--QVLI 432
Db 292 -----LDWRDVKGVSGWGLLLFGGIALSGIMKKTGTAKFISQELVDLH 337
Qy 433 GLPNIVQSITFVLANFSAFNANVAVANIVLPILCEMSLALELHPLILTPACLGISMV 492
Db 338 GLPTFLFLITVLVFIPLTELMSNTATTALTAFLPILFSTAQMIGKPEMLVIPAAVAASCA 397
Qy 493 YFLPVSTPPNAIVTQYAHIKTKYFACCGIVPTIIGISVAL 532
Db 398 FMLPVATPPNAIVGTGVYKQSQMMRVGLILNIV-FSIVL 436
RESULT 14
I64080
probable membrane protein HI0608 - Haemophilus influenzae (strain Rd KW20)
C:Species: Haemophilus influenzae
C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 09-Jul-2004
C:Accession: I64080
R:Flaigsmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.
; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M.
Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, C.
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A:Reference number: A64000; MUID:95350630; PMID:7542800
A:Accession: I64080
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-461 <TIGR>
A:Cross-references: UNIPROT:Q57486; GB:U32743; GB:L42023; NID:g1573597; PIDN:AAC22267.1;
C:Superfamily: probable transporter MJ0672
Query Match 15.1%; Score 436; DB 2; Length 461;
Best Local Similarity 24.8%; Pred. No. 9.3e-26;
Matches 123; Conservative 100; Mismatches 174; Indels 98; Gaps 15;
Qy 50 LMLIVTMALLMITETPLIYVVTALFPLVFCPLLGLVNASIVCKYQYFTDTIVVFLGGLI 109
Db 40 LALLAFTAVLWLSALHTVITALLVPLLAVALGLVSTKQALVGFPADPTIPLFFGGSFLAT 99
Qy 110 GIEYSNLHTRIALRVIRIVGSSPRRLFVG---LMSVTFMGLWISNAGTAMMCPIVKAL 166
Db 100 ALHIQKLDKLIANKIMALARGN---LFIATVILPLITAFLSMMNSNTATAMMLPLAMGI 156
Qy 167 VNELDTNKIPVYMQEPEEPVEGEPPHPSKITVAFYAGIAYASSIGGLTGLTGTNLV 226
Db 157 LSQDREKDNHTY-----VFVLGIAYSASIGMGTLVSGPPNAI 196
Qy 227 FRGIYTERPPTSTVEITPANFMYSPILMVIWVTLVIAFLITHMGLFRPNSKTGKIIA 286
Db 197 V-----ASNLLTFSDWLWGLPIMILLPLMIGILVII-----FRPK-----L 235
Qy 287 EANTNRKLMEDVLRQRHIDLGPMSCHEIQMAIAFAPMVLITRKPGFVPGWSDLINRKY 346
Db 236 HLNPEQTF-----ENIENPM-----RLTPIIFPIALT-----W--IFSOKI 272
Qy 347 VGSASGL-----SF--VLILFALPTQYTFPKYCCGKGPFTAQAIADAILSWEVLRN 396
Db 273 NPFISGLGLQKNIASPDSIVALLAAI-----VICSTG-----VASWKIQSN 315

Qy	397	IPNGLIFLGGGFALAVASRETGINIMISKAMQVLI-GLPNIVVQSIITFVLNFFSAFNA	455	
Db	316	TDWGVLMFGGGLTSAVLKDSGASILADSIYFMDGGHXYLIGLVLVAAPFIIFLTFEPTS	375	
Qy	456	NVVAVANIVLPILCEMSLALHPLIITLPACLGISVMVYFLPVPSTPPNAIVTQVAHIKTKY	515	
Db	376	NTASALLVPIFISIAQSLGMPGILGALLIGIGASCAPMLPVATPPNAIVFGSGQVKQSE	435	
Qy	516	FACCGIVPTIIGISV	530	
Db	436	MVKVGFLLNLVCVVV	450	
RESULT 15				
B82510				
transporter, NadC family VCA0025 [imported] - Vibrio cholerae (strain N16961 serogroup O				
C:Species: Vibrio cholerae				
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004				
C:Accession: B82510				
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;				
Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.				
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.				
Nature 406, 477-483, 2000				
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.				
A:Reference number: A82035; MUID:20406833; PMID:10952301				
A:Accession: B82510				
A>Status: Preliminary				
A:Molecule type: DNA				
A:Residues: 1-462 <HEI>				
A:Cross-references: UNIPROT:Q9KNE0; GB:AB004346; GB:AB003853; NID:G9657401; PIDN:AAF9593				
A:Experimental source: serogroup O1; strain N16961; biotype El Tor				
C:Genetics:				
A:Gene: VCA0025				
A:Map position: 2				
C:Superfamily: probable transporter MJ0672				
Query Match 14.9%; Score 433; DB 2; Length 462;				
Best Local Similarity 23.8%; Pred. No. 1.6e-25;				
Matches 124; Conservative 96; Mismatches 170; Indels 132; Gaps 14;				
Qy	43	DMAEFKCLW-----LIVTWALLWITETLPYVYTALPPLVFCPLGLVNASI	88	
Db	28	DVALFLALYHLPPEHNVLGISMFLAFIAVLMLTEALHVTITAILVPVMAVFFGIFETQA	87	
Qy	89	VCKQYFTDTIVVFLGGLIVALGIEVSNLHTRIALRIVIGGSPRRLFVGLMSVSTPMGL	148	
Db	88	ALNNFANSIIFLELGGFALAAAMHHQGLDKVIADKVLAMAGKMSVAVFMLFGVTALLSM	147	
Qy	149	WISNSAGTAMCPYKALVNELDNTNKIPPVYMTQEBPVEGEPPHPSKITVAFYAGIAY	208	
Db	148	WISNTATAAMMLPLVLGVLKVDADKQSTY-----VFVLLGVAY	187	
Qy	209	ASSIGGLGTLLGTGNLVFRGIYTERPFTSTVEITFANFMEYSIP-LMVIIVNVTLVIIAF	267	
Db	188	SASIGGIATLVGSPENAI-----AAAEVGLSFTDMKFKGLFTAMMMLPMAIALIYF	238	
Qy	268	LI--THMGLER-----PNSKTGKIABEANTNRKLMEDVLRQRHIDLGPMSCHETQMAIAF	320	
Db	239	LLKPTLNGMFELDRAPVNDKQKVT-----LG-----IF	268	
Qy	321	AFMIVLLITRKFGFVPGNSDLNKRKVGSA-----GLSFVLLIFALPQYTFKYCCG	375	
Db	269	GLTVFLWIFSSP-----INAAAGGFKSPDTLVALGAILMLSA-----	306	
Qy	376	KGPFTAQAIDALLSWEYVLRNIPWGLLELGGGFALAVASRETGINIMISKAMQVLI---	432	
Db	307	-----RVVHWKEIQKTADWGVLLELFGGGLCLSNVLKQGTGTSVFLANALSDMVSHM	356	
Qy	433	GLPNIVVQSIITFVLNFFSAFNAVIVNIVLPILCEMSLALHPLIITLPACLGISMV	492	
Db	357	GIFVILVATFVV--FLTEFASNTASAALLIPVATVAEAFGMSFVLLSVLIAVAASCA	414	
Qy	493	YFLPVPSTPPNAIVTQVAHIKTK-----YFACCGIVPTI	525	

Db

415 FMLPVATPPNAIVPASGHIKQSEMMRVGLYLNIAICIGLLTAI

456

Search completed: June 30, 2005, 09:00:15

Job time : 18.7848 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2005, 08:28:39 ; Search time 75.6279 Seconds
(without alignments)
3805.323 Million cell updates/sec

Title: US-10-017-479A-3

Perfect score: 2897

Sequence: 1 MAEPGEQRKFLVGRCCIFHW.....PPESKSPDWAKEIKNQTKI 562

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : UniProt_03.*

1: uniprot_eprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2897	100.0	562	1	IND2_DROME
2	1543	53.3	505	2	Q811B6
3	1494.5	51.6	504	2	Q86B89
4	1382.5	47.7	572	1	IND1_DROME
5	1382.5	47.7	590	2	Q7KUS6
6	1040	35.9	896	2	Q7QIT2
7	890	30.7	605	2	Q866R0
8	880.5	30.4	586	2	Q6AZR9
9	875	30.2	568	2	Q86YT5
10	871	30.1	568	2	Q6ZMG1
11	871	30.1	587	2	Q35055
12	868.5	30.0	586	1	S132_MOUSE
13	855.5	29.5	572	2	Q8CJ44
14	852.5	29.4	622	2	O57661
15	851.5	29.4	624	2	Q642N5
16	846	29.2	613	2	Q803K7
17	844	29.1	592	1	S132_HUMAN
18	843	29.1	602	2	Q6DJN7
19	842	29.1	613	2	Q6P312
20	840	29.0	600	1	S133_MOUSE
21	839	29.0	602	2	Q6E7G8
22	833	28.8	600	1	S133_RAT
23	832.5	28.7	602	2	Q81VB1
24	831.5	28.7	602	1	S133_HUMAN
25	830	28.7	593	1	S132_RABIT
26	818	28.2	414	2	Q7PF67
27	814.5	28.1	572	2	Q67BT3
28	808	27.9	587	1	S132_RAT
29	789	27.2	601	2	Q9W7I2
30	781	27.0	596	2	Q6NMY4
31	764	26.4	625	2	Q8BZ82

ALIGNMENTS

32	763.5	26.4	582	1	NAD1_CAEEL
33	756	26.1	583	2	Q6NY28
34	754	26.0	583	2	Q6PE27
35	741	25.6	626	2	Q8N631
36	738.5	25.5	627	1	S134_HUMAN
37	724.5	25.0	577	1	YKG6_CAEEL
38	698.5	24.1	601	2	Q7ZWL2
39	698	24.1	520	2	Q8N8K4
40	696	24.0	595	1	S131_HUMAN
41	687	23.7	595	1	S131_RAT
42	680.5	23.5	594	1	S131_MOUSE
43	680.5	23.5	594	2	Q80YB5
44	677	23.4	551	1	NAD2_CAEEL
45	676	23.3	619	2	Q6L970
46	673.5	23.2	566	1	NAD3_CAEEL
47	639	22.1	546	2	Q65NCO
48	618	21.3	520	2	Q8NV55
49	618	21.3	520	2	Q6G816
50	618	21.3	520	2	Q6GFE0
51	617	21.3	520	2	Q99SX1
52	617	21.3	520	2	Q7A4P8
53	580	20.0	495	2	Q67T81
54	572	19.7	548	2	Q6ZFH7
55	571	19.7	552	2	Q8NE33
56	551	19.0	510	2	Q8NTS7
57	551	19.0	527	2	Q6M8D0
58	545	18.8	552	2	Q25003
59	544	18.8	510	2	Q7UUK9
60	544	18.8	540	2	Q6MAW4
61	543	18.7	540	2	Q8LG88
62	540	18.6	462	2	Q9FGK1
63	536	18.5	548	2	Q9ZML8
64	533	18.4	626	2	Q6AQ77
65	528	18.2	587	2	Q8FU27
66	527	18.2	540	2	Q6FEH9
67	520.5	18.0	524	2	Q6NKS9
68	517.5	17.9	474	2	Q65MM0
69	496.5	17.1	450	2	Q75MH3
70	474	16.4	450	2	Q8FMZ5
71	463.5	16.0	302	2	Q68D44
72	459.5	15.9	461	2	Q7N822
73	447	15.4	470	2	Q65R88
74	446.5	15.4	474	2	Q8GF70
75	444.5	15.3	456	2	Q7MLG6
76	444.5	15.3	456	2	Q8D908
77	442.5	15.3	463	2	Q9CLO0
78	442	15.3	445	2	Q66449
79	436	15.1	461	1	Y608_HAEIN
80	433	14.9	462	2	Q9KNE0
81	428.5	14.8	459	2	Q66615
82	428.5	14.8	459	2	Q8C2T3
83	427.5	14.8	456	2	Q6LQW4
84	425	14.7	456	2	Q87Q94
85	425	14.7	471	2	Q9JUV43
86	425	14.7	471	2	Q9K032
87	420.5	14.5	456	2	Q8ZHX0
88	418	14.4	464	2	Q8VTX1
89	418	14.4	472	2	Q8DAD2
90	416	14.4	464	2	Q9S4F3
91	415.5	14.3	463	2	Q65S85
92	415	14.3	464	2	Q8F5L4
93	415	14.3	472	2	Q7MJT9
94	414	14.3	464	2	Q9AED9
95	414	14.3	464	2	Q72QJ5
96	399	13.8	464	2	Q9ZGU3
97	396	13.7	471	2	Q6LSU9
98	396	13.7	487	2	Q9KSE3
99	388.5	13.4	451	2	Q7MBY1
100	386	13.3	491	2	Q847W4

Q93655	caenorhabdi
Q6ny28	brachydanio
Q6pe27	brachydanio
Q8n631	homo sapien
Q9ukg4	homo sapien
P45556	caenorhabdi
Q7zw12	xenopus lae
Q8n8k4	homo sapien
Q9bzv2	homo sapien
Q07782	rattus norv
Q9jhi4	mus musculu
Q80ybs	mus musculu
P32739	caenorhabdi
Q21339	anguilla ja
Q61970	caenorhabdi
Q21339	caenorhabdi
Q65nco	caecilius li
Q8nv55	staphylococ
Q6g816	staphylococ
Q6gfe0	staphylococ
Q99sx1	staphylococ
Q7a4p8	staphylococ
Q67t81	syndicobacte
Q6zfh7	oryza sativ
Q8ene3	oceanobacil
Q8nt87	corynebacte
Q6m8d0	corynebacte
Q25003	helicobacte
Q7uuk9	rhodospirell
Q9maw4	arabidopsiis
Q81g88	arabidopsiis
Q9fgk1	arabidopsiis
Q9zml8	helicobacte
Q6aq77	deaultofale
Q8fu27	corynebacte
Q9feh9	arabidopsiis
Q6nks9	corynebacte
Q65mm0	bacillus li
Q75mh3	homo sapien
Q8fmz5	corynebacte
Q68d44	homo sapien
Q7n822	photorhabdu
Q65r88	mannheimia
Q8gf70	photorhabdu
Q7mlg6	vibrio vuln
Q8d908	vibrio vuln
Q9clq0	pasteurella
Q66449	aquifex aeo
Q57486	haemophilus
Q9kne0	vibrio chol
Q66615	yersinia ps
Q8czt3	yersinia ps
Q61qm4	photobacter
Q87q94	vibrio para
Q91v43	neisseria m
Q9k032	neisseria m
Q8zhx0	yersinia pe
Q8vtx1	leptospira
Q8dad2	vibrio vuln
Q9e4f3	leptospira
Q65s85	mannheimia
Q8f5l4	leptospira
Q7mjt9	vibrio vuln
Q9aed9	leptospira
Q72qj5	leptospira
Q9zgu3	leptospira
Q6lsu9	photobacter
Q9kse3	vibrio chol
Q7m8y1	wolinella s
Q847w4	vibrio para

```

RESULT 1
ID_2 DROME STANDARD; PRT; 562 AA.
AC Q9VQDQ; Q95783;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 25-JAN-2005 (Rel. 46, Last annotation update)
DE I'm not dead yet protein 2.
GN Name=Indy-2;
OS Drosophila melanogaster (Fruit fly);
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Surtan G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Boutek J., Brokstein P., Brottier P.,
RA Borkova D., Botchan M.R., Boutek J., Brokstein P., Brottier P.,
RA Burtis K.C., Bueam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foaier C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP GENOME REANNOTATION.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.-J., Drysdale R.A.,
RA Harris N.L., Richter J.N., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review.";
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley; TISSUE=Head, and Testis;

```

```

RX MEDLINE=22426066; PubMed=12537569;
RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,
RA George R.A., Guarin H., Krommiller B., Pacleb J.M., Park S., Wan K.H.,
RA Rubin G.M., Celniker S.E.;
RT "A Drosophila full-length cDNA resource.";
RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).
RN [4]
RP IDENTIFICATION.
RX PubMed=1118146; DOI=10.1126/science.290.5499.2137;
RA Rogina B., Reenan R.A., Nilsen S.P., Helfand S.L.;
RT "Extended life-span conferred by cotransporter gene mutations in
RT Drosophila.";
RL Science 290:2137-2140(2000).
CC -!- FUNCTION: Cation-independent electroneutral transporter (not
CC associated with membrane depolarization) of a variety of
CC tricarboxylic and dicarboxylic acid-cycle intermediates. There is
CC also small, but detectable, transport of monocarboxylics.
CC Transport is through the epithelium of the gut and across the
CC plasma membranes of organs involved in intermediary metabolism and
CC storage (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -!- SIMILARITY: Belongs to the SLC13A transporter family. NADC
CC subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AE003728; AAN13819.1; -
CC EMBL; AY060287; AAL25326.1; -
CC EMBL; BT001329; AAN71084.1; -
CC FlyBase; Fgn0011603; Indy-2.
CC GO; GO:0005887; C:integral to plasma membrane; ISS.
CC GO; GO:0015137; F:cyturate transporter activity; ISS.
CC GO; GO:0050833; F:pyruvate transporter activity; ISS.
CC GO; GO:0015141; F:succinate transporter activity; ISS.
CC GO; GO:0015746; P:cyturate transport; ISS.
CC GO; GO:0008340; P:determination of adult life span; ISS.
CC GO; GO:0006848; P:pyruvate transport; ISS.
CC GO; GO:0015744; P:succinate transport; ISS.
CC InterPro; IPR001898; Na/sul symport.
CC Pfam; PF00939; Na sulph sympt. 1.
CC PROSITE; PS01271; NA_SULFATE; FALSE_NEG.
CC Transmembrane; Transmembrane.
CC TRANSMEM 24 44 Potential.
CC TRANSMEM 64 84 Potential.
CC TRANSMEM 93 113 Potential.
CC TRANSMEM 135 155 Potential.
CC TRANSMEM 200 220 Potential.
CC TRANSMEM 249 269 Potential.
CC TRANSMEM 308 328 Potential.
CC TRANSMEM 350 370 Potential.
CC TRANSMEM 393 413 Potential.
CC TRANSMEM 430 450 Potential.
CC TRANSMEM 478 498 Potential.
CC TRANSMEM 515 535 Potential.
CC SSEQUENCE 562 AA; 61934 MW; D9F04C05FAA73ED3 CRC64;
SQ

```

Query Match 100.0%; Score 2897; DB 1; Length 562;
 Best Local Similarity 100.0%; Pred. No. 5.5e-187;
 Matches 562; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 MAEPGEQRKVLGRCCIFHWRGKASIIIPITITPILYIGFQTDMAEFKCLWLIVTMALLW 60
Db 1 MAEPGEQRKVLGRCCIFHWRGKASIIIPITITPILYIGFQTDMAEFKCLWLIVTMALLW 60
Qy 61 ITETLPIYVTALPFLVFCPLGLVNASIVCKYQKFTDTITVIVFLGGLIVALGIEYSNLHTRI 120
Db 61 ITETLPIYVTALPFLVFCPLGLVNASIVCKYQKFTDTITVIVFLGGLIVALGIEYSNLHTRI 120

```

```
QY 121 ALRVIRIVGSPRRLLFVGLMSVSTFMGLWISNSAGTAMMCPVKAALVNELDNKNIPVPM 180
DB 121 ALRVIRIVGSPRRLLFVGLMSVSTFMGLWISNSAGTAMMCPVKAALVNELDNKNIPVPM 180
QY 181 TOEEEPVEGEPPHPSKITVAFVAGIAYASSIGGLTGLTGTNLVFRGIYTERPPTSTV 240
DB 181 TOEEEPVEGEPPHPSKITVAFVAGIAYASSIGGLTGLTGTNLVFRGIYTERPPTSTV 240
QY 241 EITFANFMFYSIPLMIVNVTLVIAFLITHMGLFRPNSKTGKIIAEANTNKLKLMEDVLR 300
DB 241 EITFANFMFYSIPLMIVNVTLVIAFLITHMGLFRPNSKTGKIIAEANTNKLKLMEDVLR 300
QY 301 QRHIDLGPMSCHIEIOAIAFAMIVLLITRKPGFVPGWSDLNKRKVGSGSLSFIVLLI 360
DB 301 QRHIDLGPMSCHIEIOAIAFAMIVLLITRKPGFVPGWSDLNKRKVGSGSLSFIVLLI 360
QY 361 FALPTQYTFEKKYCCGKGPPTAQDAIDALSWEYVLRNIPMGLLFLGGGFPALAVASRETGL 420
DB 361 FALPTQYTFEKKYCCGKGPPTAQDAIDALSWEYVLRNIPMGLLFLGGGFPALAVASRETGL 420
QY 421 NIMISKAMQVLGLNIVVQSITFVLNFFSAFNANVVANIVLPLICMSLALHPLI 480
DB 421 NIMISKAMQVLGLNIVVQSITFVLNFFSAFNANVVANIVLPLICMSLALHPLI 480
QY 481 LTLPACLGISMYVFLPVSTPPNAIVTQYAHIKTKYPACCGIYPTIIGISVALVNTNTWGL 540
DB 481 LTLPACLGISMYVFLPVSTPPNAIVTQYAHIKTKYPACCGIYPTIIGISVALVNTNTWGL 540
QY 541 IIPPEKSPDPAWKEIKNQTKI 562
DB 541 IIPPEKSPDPAWKEIKNQTKI 562

RESULT 2
QB11B6 PRELIMINARY; PRT; 505 AA.
AC Q811B6
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Indy-2-PA.
GN Name=Indy-2;
OS Drosophila pseudoobscura (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7237;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tucson 14011-0121.4;
RA Bergman C.M., Pfeiffer B.D., Rincon-Limas D.E., Hoskins R.A., Park S.,
RA Gnietke A., Mungall C.J., Wang A.M., Krommiller B., Pacleb J., Rubin G.M.,
RA Stapleton M., Wan K., George R.A., de Jong P.J., Botas J., Rubin G.M.,
RA Celniker S.E.;
RT Assessing the impact of comparative genomic sequence data on the
RT functional annotation of the Drosophila genome."
RL Genome Biol. 3:research0086-research0086(2002).
DR EMBL, AY199949; AAO01056.1; -
DR FlyBase; FBgn0064420; DpseIndy-2.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006814; P:sodium ion transport; IEA.
DR InterPro; IPR001395; Aldo/ket_red.
DR InterPro; IPR001898; Na/sul_symp.
DR Pfam; PF00939; Na_sulph_symp; 2.
DR PROSITE; PS00063; ALDOXETO_REDUCTASE_3; UNKNOWN 1.
SQ SEQUENCE 505 AA; 55608 MW; C1F8E3D37C579A40 CRC64;
```

```
Query Match 53.3%; Score 1543; DB 2; Length 505;
Best Local Similarity 56.3%; Pred. No. 7.8e-96;
Matches 285; Conservative 96; Mismatches 123; Indels 2; Gaps 2;
```

```
QY 50 LMLIVTMALLMITETLPIYVTALPPLVFCPLGLVNASIVCKQYQYTDITVIVPLGLLVAL 109
DB 50 LMLIVTMALLMITETLPIYVTALPPLVFCPLGLVNASIVCKQYQYTDITVIVPLGLLVAL 109
QY 1 MYLVSNLFWITTEALPLYSLSFPVFLPLFDILGSDQVCKLYFSDTVVMFVIGGLIIL 60
DB 1 MYLVSNLFWITTEALPLYSLSFPVFLPLFDILGSDQVCKLYFSDTVVMFVIGGLIIL 60
QY 110 GIEYSLNLTALRIVIRIVGSPRRLLFVGLMSVSTFMGLWISNSAGTAMMCPVKAALVNE 169
DB 110 GIEYSLNLTALRIVIRIVGSPRRLLFVGLMSVSTFMGLWISNSAGTAMMCPVKAALVNE 169
QY 61 AIEYSLNLTALRIVIRIVGSPRRLLFVGLMSVSTFMGLWISNSAGTAMMCPVKAALVNE 120
DB 61 AIEYSLNLTALRIVIRIVGSPRRLLFVGLMSVSTFMGLWISNSAGTAMMCPVKAALVNE 120
QY 170 LDTNKFVPMYVTOEEEPVEGEPPHPSKITVAFVAGIAYASSIGGLTGLTGTNLVFRG 229
DB 170 LDTNKFVPMYVTOEEEPVEGEPPHPSKITVAFVAGIAYASSIGGLTGLTGTNLVFRG 229
QY 121 LDSQNIPTVYKSQEEEPMEEGDPHPSTISMAFYGVAYATIGCGTGLTGTNLTPKG 180
DB 121 LDSQNIPTVYKSQEEEPMEEGDPHPSTISMAFYGVAYATIGCGTGLTGTNLTPKG 180
QY 230 IYTERPPTSTVEITFANFMFYSIPLMIVNVTLVIAFLITHMGLFRPNSKTGKIIAEAN 289
DB 230 IYTERPPTSTVEITFANFMFYSIPLMIVNVTLVIAFLITHMGLFRPNSKTGKIIAEAN 289
QY 181 LYDTRFPKSKTQIDPFIPMAYAIPTVIVNVLLYFSLQVTHMGLFRGKTSQGLEVKGRT 240
DB 181 LYDTRFPKSKTQIDPFIPMAYAIPTVIVNVLLYFSLQVTHMGLFRGKTSQGLEVKGRT 240
QY 290 TNRKLMEDVLRORHIDILGPMSCHIEIOAIAFAMIVLLITRKPGFVPGWSDLNKRKVG 349
DB 290 TNRKLMEDVLRORHIDILGPMSCHIEIOAIAFAMIVLLITRKPGFVPGWSDLNKRKVG 349
QY 241 EGQAVVKTIVIKARHQLGPGMTCHIEIQVTLVILMVFLLFTRKPGFVPGWSDLNKRKVG 300
DB 241 EGQAVVKTIVIKARHQLGPGMTCHIEIQVTLVILMVFLLFTRKPGFVPGWSDLNKRKVG 300
QY 350 ASGLSFIVLILFALPTQYTFEKKYCCGKGPPTAQDAIDALSWEYVLRNIPMGLLFLGGG 409
DB 350 ASGLSFIVLILFALPTQYTFEKKYCCGKGPPTAQDAIDALSWEYVLRNIPMGLLFLGGG 409
QY 301 GPPVWLPMVLLFALPTQYTFEKKYCCGKGPPTAQDAIDALSWEYVLRNIPMGLLFLGGG 359
DB 301 GPPVWLPMVLLFALPTQYTFEKKYCCGKGPPTAQDAIDALSWEYVLRNIPMGLLFLGGG 359
QY 410 ALAVASRETGLNIMISKAMQVLGLNIVVQSITFVLNFFSAFNANVVANIVLPLICE 469
DB 410 ALAVASRETGLNIMISKAMQVLGLNIVVQSITFVLNFFSAFNANVVANIVLPLICE 469
QY 360 ALAEGSKVSGMAKMLGDSLKPAKMPPIVVEGMCILIGLFCATFSSNAICNILIPSE 419
DB 360 ALAEGSKVSGMAKMLGDSLKPAKMPPIVVEGMCILIGLFCATFSSNAICNILIPSE 419
QY 470 MSLALELHPLITLPACLGISMYVFLPVSTPPNAIVTQYAHIKTKYPACCGIYPTIIGIS 529
DB 470 MSLALELHPLITLPACLGISMYVFLPVSTPPNAIVTQYAHIKTKYPACCGIYPTIIGIS 529
QY 420 MALAIKHPKLTLPPLPSLAISMAVHLPVSTPPNAIISGYAGIKTKYLALAGILPTIWA 479
DB 420 MALAIKHPKLTLPPLPSLAISMAVHLPVSTPPNAIISGYAGIKTKYLALAGILPTIWA 479
QY 530 VALVNTNTWGLIIPPEKSPDPAWKEIKNQTKI 555
DB 530 VALVNTNTWGLIIPPEKSPDPAWKEIKNQTKI 555
QY 480 VLMLNSQ-YGNIIPPEKSPDPAWKEIKNQTKI 504
DB 480 VLMLNSQ-YGNIIPPEKSPDPAWKEIKNQTKI 504

RESULT 3
QB6B89 PRELIMINARY; PRT; 504 AA.
AC Q86B89
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE CG32921-PB (CG32921-pd).
GN ORFNames=CG32921;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.B.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Frannkoch C., Baldwin D.,
RA Balles R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.V., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Dou L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwan C.,
```

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Neelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodruff, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
RN [2]
RN SEQUENCE FROM N.A.
RP MEDLINE=22426065; PubMed=12537568;
RX Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Swirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: Release 3 of the *Drosophila*
RT melanogaster euchromatic genome sequence.";
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [3]
RN SEQUENCE FROM N.A.
RP MEDLINE=22426070; PubMed=12537573;
RX Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Swirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the *Drosophila melanogaster* euchromatin:
RT a genomic perspective.";
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
RN [4]
RN SEQUENCE FROM N.A.
RP MEDLINE=22426069; PubMed=12537572;
RX Mira S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a
RT systematic review.";
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RN [5]
RN SEQUENCE FROM N.A.
RP FlyBase;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003728; AAC041577.1; -;
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006814; P:sodium ion transport; IEA.
DR InterPro; IPR001898; Na/sulph_symp; 2.
DR Pfam; PF00939; Na sulph. sym; 2.
SQ SEQUENCE 504 AA; 55386 MW; 28FDBFCFE5065F4A CRC64;
Query Match 51.6%; Score 1494.5; DB 2; Length 504;
Best Local Similarity 54.2%; Pred. No. 1.4e-92;
Matches 273; Conservative 94; Mismatches 136; Indels 1; Gaps 1;

QY 50 LMLVITMALLWITETLPIYVTALPFLVFCPLGLGLVNASIVCKQYFTDIYVFLGGLIAL 109
Db 1 MYLVANMALFWITTEALPIYLTALFPVFLPLFGILTSEKVCFSYFSDTWVWFIIGLLIAL 60
QY 110 GIEYSNLHTRIALRIVIRVIGSGPRRLFVGLMSVSTFMGLWINSAGTAMCMCIVKALVNE 169
Db 61 AIEYSNLHQRIALNTILIVGCSPRRLHFLGVMTCTFISLWINSAAATAMCMCIVKALVNE 120
QY 170 LDTNKIFPVYMTQEEPEVEEGPPHPSKITVAFYAGIAYASSIGLGLTGLTGNTLVFRG 229
Db 121 METQNIFFAIYKTOEEPEVEEGPPHPSKITVAFYAGIAYASSIGLGLTGLTGNTLVFRG 180
QY 230 IYTERFPTSTVETITFANFMFYISPLMVIIVNVLVIIAFILITHMGLFRPNSTKTKIIAEAN 289
Db 181 LYDTRFNSDEKIDPIFMAYSVFVFLVLIELTYFSLQVTHMGLFRPNSTKTKIQQEVKGA 240
QY 290 TRKLMEDVLRQHDLDGMSCHETQMAIAPAFMIVLLITRKPGVPGVGSDLINRKVCS 349
Db 241 ESQDVVKDKVQKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEK 300
QY 350 ASGLSPVILLIEALPTQYTFYKCGKGPFTAQAIDAILSWBYVLRNIPWGLLFLGGGF 409
Db 301 GPPVFAITILLFALPTQYTFYKCGKGPFTAQAIDAILSWBYVLRNIPWGLLFLGGGF 359
QY 410 ALAVASRETGLNIMISKAMQVILGPNVIVQSIITFVLNFFAFNANVNVVIVLPILCE 469
Db 360 ALAEGSRVSGMAKLGESLAFAGEMHSLVLSICIIISLCTAFASNAICNLIPIFSE 419
QY 470 MSALALEHPLIITLPACLGISVMYFELPVSTPNAIVTQVAHITKYFACCGIVPTIIGIS 529
Db 420 MALAIEVHPMKUTFPALACSLAFHLPVSTPNAIVTQVAHITKYFACCGIVPTIIGIS 479
QY 530 VALVNTNTWGLIIFPESKSPDWA 553
Db 480 CLLFTGVTMTLIYPTGTFEPPSWA 503
RESULT 4
INDI DROME STANDARD; PRT; 572 AA.
AC Q9VWT2; Q9NHV9;
DT 05-JUL-2004 (Rel. 44, Created)
DT 25-JUL-2004 (Rel. 44, Last sequence update)
DT 25-JAN-2005 (Rel. 46, Last annotation update)
DE I'm not dead yet protein (INDY transporter protein) (drIndy).
GN Name=Indy; ORFNames=CG3979;
OS *Drosophila melanogaster* (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; *Drosophila*.
OC NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A., AND FUNCTION.
RX MEDLINE=2250702; PubMed=12186628; DOI=10.1042/BJ20021132;
RA Inoue K., Fei Y.-J., Huang W., Zhuang L., Chen Z., Ganapathy V.;
RT "Functional identity of *Drosophila melanogaster* Indy as a cation-
RT independent, electroneutral transporter for tricarboxylic acid-cycle
RT intermediates.";
RL Biochem. J. 367:313-319(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,


```
Query Match      30.7%; Score 890; DB 2; Length 605;
Best Local Similarity 34.9%; Pred. No. 8.9e-52;
Matches 204; Conservative 105; Mismatches 207; Indels 68; Gaps 12;

QY 25 SIILPLITPLIYGFOTDMAEFKCLMLVIMALLWITETLPIYVVTALFPLVFCPLGLGLV 84
DB 17 SILFILLPLPLL---VPTKEAKCAVSIIIMALLWCTETLPLAVTAFPLIIFPPMMGIM 73

QY 85 NASIVCKQYFTDTIVFLGGLIIVAGLHGYEYSLHTRIALRVIRVGGSPRFLFGLMSVST 144
DB 74 DASEVSEYKLDITNIDFGLGLVIAVEHWNHKLRIALRVLLIIGVRPAFLIIGLGMVITA 133

QY 145 FMGLWINSAGTAMCPIVAKVNL-----DT----- 172
DB 134 FLSMWISNTATTAMVPIAHVLSQLHKGPKEKDTMGHVNISFELQEPHGNPKKEPSSL 193

QY 173 -----NKIPVYMTQBEPEVEEG-----EPHPSKITVAFYAGIYASSIGGLGTLI 219
DB 194 REKENSVPVPTSMPPPEYKEKEEKEENKEKEH-FKLSQGSMLVCVYAASIGGIATLT 252

QY 220 GTGNLVFRGIYTERPTSTVEITFANFMYSLPLMVIVNVLVIATL---ITHWGL-F 275
DB 253 GTTNLVLOGQNSLFPKNSVNFASFVGFAPFTWVL-----LILLSWIMWLQILFLGNF 308

QY 276 RPSKTKGIIAEANTNKLMDVLQRHIDLGPMSCHIEOMIAFAPMIVLLITRKPGFV 335
DB 309 RKNFNCGR---KSKQEKRAAYQVLOTBHKLGPMSFAEIAVTFLFLLVNLVTRFPGFF 365

QY 336 PGWSDLI-----NRKVVGSAGSLFVLLIFALPTQVTFPKYCGKGPFTQAIDAILSW 390
DB 366 PGWGNVAFNSDEGSMVSDGVAFVFIILFIVPSLVPFSQQGSKF--KAPPALLDW 423

QY 391 EYVLNIPWGLFLGGLFALAVASRETGLNIMISKAMOVILGLPNIVQSIITFVLNPF 450
DB 424 NTVNKKQVPMNIFLGLGGFALAKGSEVSLGSLMGNKLTPLQSIPIAPAFILCLLVATF 483

QY 451 SAFVANVAVNIPLVICENSALDELHPLILTLPACIGISWYVPLPVSTPPNAIVTQYAH 510
DB 484 TECASNATVTLFPLILASMAQAICNPLVLMPECTLSIASLAPMLVATPPNAIVSYGQ 543

QY 511 IKTKYFACCGIVPTIIGISVALVNTNTWGLIIFPESKSPFDWAK 554
DB 544 LKVIDMAKTGFLNIIGVLTITLAINWSPYIF-QLDQDFTWAQ 586

RESULT 8
ID Q6AZR9 PRELIMINARY; PRT; 586 AA.
AC Q6AZR9;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Slc13a2-prov protein.
OS Name=slc13a2-prov;
GN Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RL Dev. Dyn. 225:384-391 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX PubMed=12477932; DOI=10.1073/pnas.242603899;

Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner R., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Villalón D.K., Munz D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
Blakesley R.C., Touchman J.W., Green E.D., Dickson M.C.,
Rodriguez A.C., Guimond J., Schmitz J., Myers J.M., Butterfield Y.S.,
Krzywinski M.I., Skalek U., Smailus D.E., Schnerch A., Schein J.E.,
Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Klein S., Gerhard D.S.;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
EMBL: BC077435; AAH77435.1; -.
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0005215; F:transporter activity; IEA.
DR GO: GO:0006814; P:sodium ion transport; IEA.
DR InterPro: IPR001898; Na/sulph symport.
DR Pfam: PF00939; Na sulph sym; 1.
DR PROSITE: PS01271; NA SULFATE; 1.
SQ SEQUENCE 586 AA; 64959 MW; 2E3646A1D81F31DB CRC64;

Query Match      30.4%; Score 880.5; DB 2; Length 586;
Best Local Similarity 35.2%; Pred. No. 3.8e-51;
Matches 201; Conservative 101; Mismatches 210; Indels 59; Gaps 9;

QY 30 LITPLIYGFOTDMAEFKCLMLVIMALLWITETLPIYVVTALFPLVFCPLGLVNASIV 89
DB 24 LPLPLLI---QT---KEASCAVILMAYVWCTEVIPLAVTALLPVLFPFMGILESKV 78

QY 90 CKQYFTDTIVFLGGLIIVAGLHGYEYSLHTRIALRVIRVGGSPRRLFVGLMSVSTPMGLW 149
DB 79 CMQYLKDTNMLFVGLIIVAVAVEQNLHKRIALKVLLIVGVRPALMLGLFMGVTAFLSMW 138

QY 150 INSAGTAMCPIVAKVNL-----ELDTNKLFPVYMTQ---- 182
DB 139 ISNTATTAMVPIVQAVLSQLHTAEEDPSMLSELEGQTNPALETKNAIPMQLVQTVSN 198

QY 183 -----EESPVEEGEPHPSPKITVAFYAGIYASSIGGLGTIGTGNLVFRGIYTERFP 236
DB 199 GHVPEVPEVPEVPEVPEVPEVPEVPEVPEVPEVPEVPEVPEVPEVPEVPEVPEVPEV 255

QY 237 TSTVEITFANFMYSIPLMVIVNVLVIATLTHMGLFRPNSTKGIKIIAEANTNKLME 296
DB 256 NNGDILNFAWFGFAFPNNVIM---LCLSWFLQPSFIGNFKKTGWCGCATASEKRAAY 312

QY 297 DVLRQRHIDLGPMSCHIEOMIAFAPMIVLLITRKPGFVPGWSDLI-----NRKVVGSAG 352
DB 313 SVIREEYRKLGPISVAESVLFLLFLLVLLWFTDRDPGVVGNATILFNKDNIEYATDATV 372

QY 353 LSFVILLIFALPTQYTFYKVCCKGPGFTAQID-----AILLSWYVLNRNIPWGLL 402
DB 373 AVFVAFLLFILPATPKFCCCKARNSFDEDEBEQKDTFFSAPILTWKVVQKQWPSIV 432

QY 403 FLGGGFALAVASRETGLNIMISKAMQVILGLPNIVQSIITFVLNFFAFNANVVAVNI 462
DB 433 LLLGGGFALAKGSDASGLSHWLGGQWTPHLSIPPPWAIILSLMIAVFTFECASNATATL 492

QY 463 VLPILCEMSLALHPLIILTLFACLGISWYVPLPVSTPPNAIVTQYAHILTKYFACCGIV 522
DB 493 FLPLILASMSRSIEVNPLYIMIPCTLTSTSPAFMLPVATPPNAIVFVSGHLRVRSDMVKTV 552
```


Query Match 30.0%; Score 868.5; DB 1; Length 586;
Best Local Similarity 36.6%; Pred. NO. 2.4e-50;

Search completed: June 30, 2005, 08:58:43
Job time : 77.6279 secs

THIS PAGE IS BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2005, 08:28:14 ; Search time 82.8208 Seconds
(without alignments)
2741.199 Million cell updates/sec

Title: US-10-017-479a-4

Perfect score: 3058

Sequence: 1 MATCPALWAYRYLYLVLC.....HSNTSQCLLNPSNVTVPGL 587

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : A_Geneseq_16Dec04.*

1: Geneseqp1980a.*

2: Geneseqp1990a.*

3: Geneseqp2000a.*

4: Geneseqp2001a.*

5: Geneseqp2002a.*

6: Geneseqp2003a.*

7: Geneseqp2003bs.*

8: Geneseqp2004a.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	3058	100.0	587	7	AAE38766	Aae38766	Rat 69624
2	2392	78.2	592	6	ABB82950	Abb82950	Human SLC
3	2392	78.2	592	6	ABB82951	Abb82951	Human SLC
4	2392	78.2	592	6	ABR40100	AbR40100	Human ren
5	2392	78.2	592	6	ABR57025	AbR57025	Human NaD
6	2392	78.2	592	7	AD663755	Ad663755	Human Pro
7	2073.5	67.8	519	8	ABM84184	Abm84184	Human tra
8	2064.5	67.5	519	8	ADH22543	Adh22543	Human tra
9	1971.5	64.5	619	8	ADG16995	Adg16995	African c
10	1829	59.8	581	8	ADP64799	Adp64799	Zebrafish
11	1828	53.2	616	5	ABU65064	Abu65064	Human NOV
12	1628	53.2	616	5	ADU51048	Adu51048	Human NOV
13	1628	53.2	616	7	ADK51052	Adk51052	Human NOV
14	1628	53.2	616	8	ADH42447	Adh42447	Human NOV
15	1628	53.2	616	8	ADN61777	Adn61777	Human NOV
16	1628	53.2	616	8	ADN61779	Adn61779	Human NOV
17	1624	53.1	568	8	ABR57023	AbR57023	Human TCH
18	1624	53.1	568	7	AAE38764	Aae38764	Human 696
19	1624	53.1	568	7	ADK51052	Adk51052	Human NOV
20	1624	53.1	568	8	ADH42441	Adh42441	Human NOV
21	1624	53.1	568	8	ADP64793	Adp64793	Human Na+
22	1624	53.1	568	8	ADH42443	Adh42443	Human NOV
23	1622	53.0	568	5	ABU65062	Abu65062	Human NOV
24	1622	53.0	568	7	ADK51050	Adk51050	Human NOV
25	1622	53.0	568	8	ADH42445	Adh42445	Human NOV

99 773 25.3 510 4 AAG90000 Aag90000 C glutami
100 706.5 23.1 304 5 ABB89646 ABB89646 Human pol

ALIGNMENTS

RESULT 1
AAE38766
ID AAE38766 standard; protein; 587 AA.
XX AAE38766;

XX 18-DEC-2003 (first entry)

XX Rat 69624 protein sodium sulphate symporter domain.

XX Rat; 69624; transporter protein; neurological disorder; therapy;
XX atherosclerosis; cardiac hypertrophy; ischaemia reperfusion injury;
XX metabolic disorder; haematopoietic neoplastic disorder; leukaemia;
XX arthritis; multiple sclerosis; encephalomyelitis; myasthenia gravis;
XX carcinoma; cell proliferation; autoimmune disorder; diabetes mellitus;
XX renal disorder; colon; hepatic disorder; hypocitranuria; calcium stone;
XX mental retardation; Canavan disease; differentiative disorder; sarcoma;
XX systemic lupus erythematosus; cardiovascular disorder; arteriosclerosis;
XX atrial fibrillation; forensic identification; pain.

XX Rattus norvegicus.

XX US2002193582-A1.

XX 19-DEC-2002.

XX 17-JUN-2002; 2002US-00173519.

XX 18-JUN-2001; 2001US-0298970P.

XX (MILL-) MILLENNIUM PHARM INC.

XX Curtis RAJ;

XX WPI; 2003-644620/61.

XX Novel 69624 polypeptide, a human transporter family member, useful for
XX treating disorders e.g. hypocitranuria, such as hypocitranuria, formation
XX of calcium stones, mental retardation abnormal body sulfate homeostasis.

XX Disclosure; Page 49-50; Opp; English.

XX The invention relates to 69624 polypeptide, a human transporter family
XX member and its corresponding nucleic acid. 69624 protein is useful for
XX developing novel diagnostic and therapeutic agents for 69624-mediated or
XX related disorders. 69624 protein act as therapeutic or diagnostic agents
XX for renal, neurological, colon or hepatic disorders. It act as diagnostic
XX targets and therapeutic agents for treating disorders such as
XX hypocitranuria, formation of calcium stones, mental retardation (Canavan
XX disease) or abnormal body sulphate homeostasis. 69624 protein may act as
XX diagnostic targets and therapeutic agents for controlling cellular
XX proliferative and/or differentiative disorders such as carcinoma,
XX sarcoma, metastatic disorder or haematopoietic neoplastic disorders e.g.,
XX leukaemia, immune disorders such as autoimmune disorders (diabetes
XX mellitus, arthritis), multiple sclerosis, encephalomyelitis, myasthenia
XX gravis, systemic lupus erythematosus, cardiovascular disorders such as
XX arteriosclerosis, atherosclerosis, ischaemia reperfusion injury, cardiac
XX hypertrophy, atrial fibrillation etc; and disorders involving abnormal or
XX excessive pain. 69624 sequence is useful as pharmacodynamic marker and is
XX also used in forensic identification of a biological sample. The present
XX sequence is rat sodium-dependent dicarboxylate/sulphate transporter (NaDC
XX -1) protein

XX Sequence 587 AA;

Query Match 100.0%; Score 3058; DB 7; Length 587;

Best Local Similarity 100.0%; Pred. No. 7.7e-300;
Matches 587; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATCHPALWAYRFLIVLCPIELPLPLIVOTKEAYCAYSILMALLWCTEALPLAVTA 60
DB 1 MATCHPALWAYRFLIVLCPIELPLPLIVOTKEAYCAYSILMALLWCTEALPLAVTA 60
QY 61 LFPVILFPLMGIMDASEVCIEYFKDTNILFVGLGMVAIAVEHWNHLKRALQVLLIIGVR 120
DB 61 LFPVILFPLMGIMDASEVCIEYFKDTNILFVGLGMVAIAVEHWNHLKRALQVLLIIGVR 120
QY 121 PALLLLGFMVLVTAFLSMWISNTATTAMVPIGHAVLEQLQSGKQVGGNNNPTFELQEE 180
DB 121 PALLLLGFMVLVTAFLSMWISNTATTAMVPIGHAVLEQLQSGKQVGGNNNPTFELQEE 180
QY 181 CPQKEVTKLDNGQPVASPEPRTKTQEHHRFSQGLSLCICYSASIGGTATLTGTTPNLV 240
DB 181 CPQKEVTKLDNGQPVASPEPRTKTQEHHRFSQGLSLCICYSASIGGTATLTGTTPNLV 240
QY 241 LOGQVNSLFPQNGVNVNFASFQGFAPPTMIILLLLAWLMLQVLFLGVNFRKNFGFGE 300
DB 241 LOGQVNSLFPQNGVNVNFASFQGFAPPTMIILLLLAWLMLQVLFLGVNFRKNFGFGE 300
QY 301 ERKQAAFOVIKTQYRLLGPMSEFAEKTVTVLVLLVLFVLFVLFVLFVLFVLFVLFV 360
DB 301 ERKQAAFOVIKTQYRLLGPMSEFAEKTVTVLVLLVLFVLFVLFVLFVLFVLFVLFV 360
QY 361 SMASDGTVAIFISLVNFIIPSKIPLGMQDPKPKGLKAPPAILTWKTVDKMPNIVILL 420
DB 361 SMASDGTVAIFISLVNFIIPSKIPLGMQDPKPKGLKAPPAILTWKTVDKMPNIVILL 420
QY 421 GGGFALAKGSEQSLSEWLGDKLTPLQHIPPSATAVILCLLIAIETECTSNVATTLFLP 480
DB 421 GGGFALAKGSEQSLSEWLGDKLTPLQHIPPSATAVILCLLIAIETECTSNVATTLFLP 480
QY 481 ILASMAQAICLHPLYVMLPCTLAASLAFMLPVATPNAIVFSGGLKVSMDARAGFLNI 540
DB 481 ILASMAQAICLHPLYVMLPCTLAASLAFMLPVATPNAIVFSGGLKVSMDARAGFLNI 540
QY 541 IGVLAITLSINSWSIPFKLDTFPPSWAHSNTSOCLLNPSNSTVPGGL 587
DB 541 IGVLAITLSINSWSIPFKLDTFPPSWAHSNTSOCLLNPSNSTVPGGL 587

RESULT 2

ABB82950
ID ABB82950 standard; protein; 592 AA.

XX ABB82950;

XX 14-APR-2003 (first entry)

XX Human SLC13A related protein (GenBank Identifier No. GI#2499523).
DE SLC13A; p53; sodium-sulfate cotransporter 2; cytosolic; cancer;
KW transmembrane protein; human.

XX Homo sapiens.

XX WO200298468-A1.

XX 12-DEC-2002.

XX 03-JUN-2002; 2002WO-US017460.

XX 05-JUN-2001; 2001US-0296076P.

XX 10-OCT-2001; 2001US-0328605P.

XX 15-FEB-2002; 2002US-0357253P.

XX 01-MAR-2002; 2002US-0361196P.

XX (EXEL-) EXELIXIS INC.

XX Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;

PI Lioubin MN;
DR WPI; 2003-167297/16.
XX
XX
XX Identifying candidate p53 pathway modulating agent for diagnosing or
PT treating cancer comprises detecting test agent-biased activity of an
PT assay system comprising purified Sodium sulfate cotransporter 2 (SLC13A)
PT polypeptide.
XX
XX
XX Claim 13; Page 59-61; 69pp; English.
XX
XX The invention relates to identifying a candidate p53 pathway modulating
CC agent. The method involves assaying purified Sodium-sulfate cotransporter
CC 2 (SLC13A) polypeptide or nucleic acid or a functionally active fragment
CC or derivative with a test agent. The methods are useful for identifying a
CC candidate p53 pathway modulating agent, modulating a p53 pathway of a
CC cell, or a mammalian cell and for diagnosing a disease in a patient. The
CC disease is breast, colon, lung or ovary cancer having greater than 25%
CC expression level. The method is useful for manufacturing a medicament for
CC diagnosing or treating breast, colon, lung or ovary cancer. Sequences
CC AB82949-952 represent polypeptide sequences related to the human SLC13A
CC protein
XX
XX Sequence 592 AA;
SQ
Query Match 78.2%; Score 2392; DB 6; Length 592;
Best Local Similarity 76.9%; Pred. No. 2e-232;
Matches 454; Conservative 63; Mismatches 67; Indels 6; Gaps 4;
Qy 1 MATCPALWAYRFLVILVCLPIFLPLPLIVQTEKAYCAYSIILMALLMCTEALPLAVTA 60
Db 1 MATCQALWAYRSYLIVFFVILLPLPLVPSKEAYCAYAIILMALPWCTEALPLAVTA 60
Qy 61 LFPVLPLMGIMDASEVCIEYFKDTNLLFVGGMLVAIAVEHNLHKKRIALQVLLIIGVR 120
Db 61 LFPVLPLMGIMDASEVAYEYKDSNLLFFGGLLVAIAVEHNLHKKRIALRVLLIIGVR 120
Qy 121 PALLLGFMLVTAFLSMWISNTATTAMVPTGHAVLEQLOGSK--KDYEGGNNPTFELQ 178
Db 121 PAPLILGFMLVTAFLSMWISNTATSMVPTAHAVLDQHSQASSNVVEGNNPTFELQ 180
Qy 179 EECPOKEVTKLDNGQ--PV-SAPSEPTQKQEHRRFSQGLSLCICYSASIGGIATLTGT 235
Db 181 EPSPOKEVTKLDNGQALPVTSSASSEGRAHLSQKHLHLTQCMSCVCYSASIGGIATLTGT 240
Qy 236 TPNVLQGVNSLFPQNGNVNFAWFGFAPPTMIILLLAWLQVLFLGVNFRKNFGF 295
Db 236 TPNVLQGVNSLFPQNGNVNFAWFGFAPPTMIILLLAWLQVLFLGVNFRKNFGF 295

XX ABB82951;
XX 14-APR-2003 (first entry)
XX Human SLC13A related protein (GenBank Identifier No. GI#4506979).
XX SLC13A; p53; sodium-sulfate cotransporter 2; cytostatic; cancer;
XX transmembrane protein; human.
XX Homo sapiens.
XX WO200298468-A1.
XX 12-DEC-2002.
XX 03-JUN-2002; 2002WO-US017460.
XX 05-JUN-2001; 2001US-0296076P.
XX 10-OCT-2001; 2001US-0328605P.
XX 15-FEB-2002; 2002US-0357253P.
XX 01-MAR-2002; 2002US-0361196P.
XX (EXEL-) EXELIXIS INC.
XX Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
PI Lioubin MN;
XX WPI; 2003-167297/16.
XX N-PSDB; ABZ24204.
XX Identifying candidate p53 pathway modulating agent for diagnosing or
PT treating cancer comprises detecting test agent-biased activity of an
PT assay system comprising purified Sodium sulfate cotransporter 2 (SLC13A)
PT polypeptide.
XX
XX Claim 13; Page 62-64; 69pp; English.
XX The invention relates to identifying a candidate p53 pathway modulating
CC agent. The method involves assaying purified Sodium-sulfate cotransporter
CC 2 (SLC13A) polypeptide or nucleic acid or a functionally active fragment
CC or derivative with a test agent. The methods are useful for identifying a
CC candidate p53 pathway modulating agent, modulating a p53 pathway of a
CC cell, or a mammalian cell and for diagnosing a disease in a patient. The
CC disease is breast, colon, lung or ovary cancer having greater than 25%
CC expression level. The method is useful for manufacturing a medicament for
CC diagnosing or treating breast, colon, lung or ovary cancer. Sequences
CC AB82949-952 represent polypeptide sequences related to the human SLC13A
CC protein
XX
XX Sequence 592 AA;
SQ
Query Match 78.2%; Score 2392; DB 6; Length 592;
Best Local Similarity 76.9%; Pred. No. 2e-232;
Matches 454; Conservative 63; Mismatches 67; Indels 6; Gaps 4;
Qy 1 MATCPALWAYRFLVILVCLPIFLPLPLIVQTEKAYCAYSIILMALLMCTEALPLAVTA 60
Db 1 MATCQALWAYRSYLIVFFVILLPLPLVPSKEAYCAYAIILMALPWCTEALPLAVTA 60
Qy 61 LFPVLPLMGIMDASEVCIEYFKDTNLLFVGGMLVAIAVEHNLHKKRIALQVLLIIGVR 120
Db 61 LFPVLPLMGIMDASEVAYEYKDSNLLFFGGLLVAIAVEHNLHKKRIALRVLLIIGVR 120
Qy 121 PALLLGFMLVTAFLSMWISNTATTAMVPTGHAVLEQLOGSK--KDYEGGNNPTFELQ 178
Db 121 PAPLILGFMLVTAFLSMWISNTATSMVPTAHAVLDQHSQASSNVVEGNNPTFELQ 180
Qy 179 EECPOKEVTKLDNGQ--PV-SAPSEPTQKQEHRRFSQGLSLCICYSASIGGIATLTGT 235
Db 181 EPSPOKEVTKLDNGQALPVTSSASSEGRAHLSQKHLHLTQCMSCVCYSASIGGIATLTGT 240
Qy 236 TPNVLQGVNSLFPQNGNVNFAWFGFAPPTMIILLLAWLQVLFLGVNFRKNFGF 295
Db 236 TPNVLQGVNSLFPQNGNVNFAWFGFAPPTMIILLLAWLQVLFLGVNFRKNFGF 295

Db 241 APNLVLQGGINSLFPQNGNVNFAWFSFAFPTMVLILLALLWLQILFLGFNFRKNFGI 300
QY 296 GEGERKQAQFQVKTQYRLLGPMSPAEKTVTVLVFLVILVLFVFTREPFGFGWDTVFA 355
Db 301 GEKMQEQQAAYCVIQTEHRLGPMTPAEKATISILFVILVLLVFTREPFGFLGWNLAFF 360
QY 356 NEKGQSMASDGTVAIFISLVMEIIPSKI PGLMODPKPKGKLKAPPAILTWKTVDNKPWN 415
Db 361 NAKGESMWSOGTVAIFIGIIMFIIPSKFPGLTQDPENFGKLKAPLGLLDWKTQVQMPWN 420
QY 416 IVILGGGFALAKGSEQSGLSWLGDKLTPLQHIPPSATAVILCLLIAIIFTECTSNVATT 475
Db 421 IVILGGGYALAKGSESGLSWLGDKLTPLQSVPAIAIILSLVATFECTSNVATT 480
QY 476 TFLPILASMAQAICLHPLYVMLPCTLAASLAFMLPVATPPNAIVSFGLKVDNARAG 535
Db 481 TIFLPILASMAQAICLHPLYVMLPCTLATSLAFMLPVATPPNAIVSFGLKVDNARAG 540
QY 536 FLNIIIGVLAITLSINSWSIPIFKLDTFSPWAHNS-TSQCLLNPSNSTVP 584
Db 541 FLNIIIGVLIILAINSWGIPFLSLHSFSPWAQSNNTAQCLPSLANTTTP 590

RESULT 4
ABR40100
ID ABR40100 standard; protein; 592 AA.
XX ABR40100;
AC ABR40100;
XX ABR40100;
DT 24-JUL-2003 (first entry)
XX Human renal sodium dicarboxylate co-transporter NADC1.
DE
XX
KW Anti-diabetic; anorectic; sodium dicarboxylate cotransporter; human;
KW sodium tricarboxylate cotransporter; HepNadc; hepatocyte; HepG2;
KW diabetes; obesity; lipid metabolism; aging; NADC1.
XX
OS Homo sapiens.
XX
XX WO2003029465-A1.
XX
PD 10-APR-2003.
XX
XX 27-SEP-2002; 2002WO-JP010038.
XX
XX 28-SEP-2001; 2001JP-00299433.
XX 28-AUG-2002; 2002JP-00249016.
XX
PA (SAKA) OTSUKA PHARM CO LTD.
XX
XX Kanemoto N, Omori Y, Sugano S, Obuchi Y;
PI WPI; 2003-354728/33.
XX
XX Sodium di- or tricarboxylate cotransporter gene (HepNadc) for treatment
PT and prevention of diabetes, obesity, for improving lipid metabolism and
PT aging.
XX
XX Example 1; Fig 1; 75pp; Japanese.
XX
XX The present invention relates to human sodium di- or tricarboxylate
CC cotransporter protein (HepNadc; ABR40097). HepNadc gene is expressed in
CC human hepatocyte cancer-origin cells HepG2. Compounds which control the
CC expression of the HepNadc gene and activity of the HepNadc polypeptide
CC may be useful for treating and preventing diabetes, obesity, for
CC improving lipid metabolism and aging. The present sequence is the NADC1
CC protein (GenBank U26209), which was used in a sequence alignment with the
CC HepNadc sequence
XX
SQ Sequence 592 AA;
Query Match 78.2%; Score 2392; DB 6; Length 592;

Best Local Similarity 76.9%; Pred. No. 26-232;
Matches 454; Conservative 63; Mismatches 67; Indels 6; Gaps 4;
QY 1 MATCHPALWAYFYLIVLCLPIFLPLPLIVOTKAYCAYSIILMALLNCTALPLAVTA 60
Db 1 MATCHQALWAYRSYITVFFVILLPLPLIVPSKEAYCAYAILMALFCTEALPLAVTA 60
QY 61 LPFIVLPFLMGIMDASEVCIEYFKDNIPLFVGLMVAIAVEHWNHKLRIALQVLLIIGVR 120
Db 61 LPFLILFPFMGIVDASEVAVEYKDSNLLFFGGLLVIAIAVEHWNHKLRIALRVLLIVGR 120
QY 121 PALLLLGFLMTAFISMWISNTATTAMVPIGHAVLEQLQSGK--KDVEGGNNNPFELQ 178
Db 121 PAPLILGFLMTAFISMWISNTATSAMVPIAHAVLDQLHSSQASSNVBEGSNNPFELQ 180
QY 179 EECPOKEVTKLNGO--PV-SAPSEPRTOKTQEHHRFSQGLSICICYSASIGGIATLTGT 235
Db 181 EFSPOKEVTKLNGQALPVTSSASSEGRAHLSQKHLHTQCMSLCVCYSASIGGIATLTGT 240
QY 236 TPNLVLQGGVNSLFPQNGNVNFAWFSFAFPTMIILLALLWLQVLFLGVNFRKNFGF 295
Db 241 APNLVLQGGINSLFPQNGNVNFAWFSFAFPTMVLILLALLWLQILFLGFNFRKNFGI 300
QY 296 GEGERKQAQFQVKTQYRLLGPMSPAEKTVTVLVFLVILVLFVFTREPFGFGWDTVFA 355
Db 301 GEKMQEQQAAYCVIQTEHRLGPMTPAEKATISILFVILVLLVFTREPFGFLGWNLAFF 360
QY 356 NEKGQSMASDGTVAIFISLVMEIIPSKI PGLMODPKPKGKLKAPPAILTWKTVDNKPWN 415
Db 361 NAKGESMWSOGTVAIFIGIIMFIIPSKFPGLTQDPENFGKLKAPLGLLDWKTQVQMPWN 420
QY 416 IVILGGGFALAKGSEQSGLSWLGDKLTPLQHIPPSATAVILCLLIAIIFTECTSNVATT 475
Db 421 IVILGGGYALAKGSESGLSWLGDKLTPLQSVPAIAIILSLVATFECTSNVATT 480
QY 476 TFLPILASMAQAICLHPLYVMLPCTLAASLAFMLPVATPPNAIVSFGLKVDNARAG 535
Db 481 TIFLPILASMAQAICLHPLYVMLPCTLATSLAFMLPVATPPNAIVSFGLKVDNARAG 540
QY 536 FLNIIIGVLAITLSINSWSIPIFKLDTFSPWAHNS-TSQCLLNPSNSTVP 584
Db 541 FLNIIIGVLIILAINSWGIPFLSLHSFSPWAQSNNTAQCLPSLANTTTP 590
RESULT 5
ABR57025
ID ABR57025 standard; protein; 592 AA.
XX ABR57025;
AC ABR57025;
XX 05-AUG-2003 (first entry)
XX Human NADC-1 amino acid sequence.
DE
XX Human; TCH169; dicarboxylate transport; hepatotrophic; cytostatic;
KW nephrotropic; vasotropic; antidiabetic; liver disease; hepatitis;
KW hepatic sclerosis; alcohol-related liver disease; prostate disease;
KW prostatitis; prostatic hypertrophy; spleen disease; spleen hyperactivity;
KW kidney disease; nephritis; kidney failure; nephritis; dropsy; diabetes;
KW diabetes-associated renal disease; metabolic disease; hyperlipaemia;
KW circulatory disease; arteriosclerosis; cancer; NADC-1.
XX
XX Homo sapiens.
XX
XX WO2003025168-A1.
PN
XX 27-MAR-2003.
PD
XX 13-SEP-2002; 2002WO-JP009444.
PF
XX 17-SEP-2001; 2001JP-00281992.
XX 02-OCT-2001; 2001JP-00306873.
PR
XX 16-APR-2002; 2002JP-00113279.
PR

XX (TAKE) TAKEDA CHEM IND LTD.
XX Nakanishi A, Uno Y, Sagiya Y;
XX WPI; 2003-313352/30.
XX Protein TCHI169 with dicarboxylate transport activity for treatment and
XX disorders.
XX Example 1; Fig 1-2; 132pp; Japanese.
XX The present invention describes protein TCHI169 and its salts having
XX dicarboxylate transport activity. TCHI169 has hepatotropic, cytostatic,
XX nephrotropic, vasotropic and antidiabetic activities. The TCHI169 protein
XX and polynucleotide can be used in the treatment, prevention and diagnosis
XX of liver disease (such as hepatitis, hepatic sclerosis and alcohol-
XX related liver disease); prostate disease (such as prostatitis and
XX prostatic hypertrophy); spleen disease (such as spleen hyperactivity);
XX kidney disease (such as nephritis, kidney failure, nephritis, dropsy and
XX diabetes-associated renal disease); metabolic disease (such as diabetes);
XX circulatory disease (such as hyperlipaemia and arteriosclerosis); and
XX cancer (such as non-small cell lung cancer, liver cancer, renal cancer,
XX ovarian cancer, prostate cancer, stomach cancer, pancreatic cancer,
XX breast cancer, colon cancer, bladder cancer and womb cancer). The present
XX sequence represents a human NADC-1 amino acid sequence which is given in
XX comparison with the human TCHI169 amino acid sequence in an example from
XX the present invention
XX Sequence 592 AA;
XX
XX Query Match 78.2%; Score 2392; DB 6; Length 592;
XX Best Local Similarity 76.9%; Pred. No. 2e-232;
XX Matches 454; Conservative 63; Mismatches 67; Indels 6; Gaps 4;
XX
Qy 1 MATCPALWAYRYLIVLCFLIFLLPLPLIVOTKEYACAYSIIMALLMCTEALPLAVTA 60
Db 1 MATCPALWAYRYLIVLCFLIFLLPLPLIVOTKEYACAYSIIMALLMCTEALPLAVTA 60
Qy 61 LPPIVLPMLGIMDAVEYVDFKDTNLLFVGLMVAIVAEHNLKRIALQVLLIGVR 120
Db 61 LPPIVLPMLGIMDAVEYVDFKDTNLLFVGLMVAIVAEHNLKRIALQVLLIGVR 120
Qy 121 PALLLGLPMLVTAFLSMVISNTATTAMVPVIGHAVLEQLQSK--KVEGGNNPTPELQ 178
Db 121 PAPILGLPMLVTAFLSMVISNTATTAMVPVIGHAVLEQLQSK--KVEGGNNPTPELQ 180
Qy 179 EECPOKEVTKLDNGQ--PV-SAPSEPRTKQTEHHRFSQGLSLCICYASIGGIATLTGT 235
Db 181 EPSPOKEVTKLDNGQALPVTSSASSEGRAHLSQKHLHLTQCMSLVCYASIGGIATLTGT 240
Qy 236 TPNVLVQGVNSLFPQNGVNVNFWGFAFPTMIILLALLMQLQVLFLGVRKQNGF 295
Db 241 APNVLVQGVNSLFPQNGVNVNFWGFAFPTMIILLALLMQLQVLFLGVRKQNGF 300
Qy 296 GEGEERKQAFQVKTQYRLLGPMSPAEKTVTLFVLLVLTFTREPGRPFGDITVEA 355
Db 301 GEKQEQQAAYCIVQTEHRLGSPMTFAKASILFVLVLLMTFTREPGFLGNGLAFF 360
Qy 356 NEKQSMASDGTVAIFISLVNFIIPSKIPGLMQDPKPKGLKAPAILTWKTVNDKMPWN 415
Db 361 NAKGESMVSDGTVAIFIGINFIIPSKIPGLMQDPKPKGLKAPAILTWKTVNDKMPWN 420
Qy 416 IVILLGGGFALAKSGESGSEWGLDKLTPLOHIPPSTAVILCLLIAIETECTSNVATT 475
Db 421 IVILLGGGFALAKSGESGSEWGLDKLTPLOHIPPSTAVILCLLIAIETECTSNVATT 480
Qy 476 TLFLPILASMAQAACLHPLYVLMCTLLASLAFMLPVATPNATVFEFGGLKVSMDARAG 535
Db 481 TIFLPIILASMAQAACLHPLYVLMCTLLASLAFMLPVATPNATVFEFGGLKVSMDARAG 540
Qy 536 FLNIIIGVLAITLSINGSWSIPFKLDTFPFSAHNSN-TSQCLLNFSNTPV 584

Db 541 FLNIIIGVLAITLSINGSWSIPFKLDTFPFSAHNSN-TSQCLLNFSNTPV 590
XX
XX RESULT 6
XX ADE63755
XX ID ADE63755 standard; protein; 592 AA.
XX AC ADE63755;
XX XX
XX DT 29-JAN-2004 (first entry)
XX DE Human Protein Q13183, SEQ ID NO 9699.
XX KW Human; pain; neuronal tissue; gene therapy;
XX KW spinal segmental nerve injury; chronic constriction injury; CCI;
XX KW spared nerve injury; SNI; Chung.
XX OS Homo sapiens.
XX XX WO2003016475-A2.
XX PN 27-FEB-2003.
XX PD 14-AUG-2002; 2002WO-US025765.
XX PF 14-AUG-2001; 2001US-0312147P.
XX PR 01-NOV-2001; 2001US-0346382P.
XX PR 26-NOV-2001; 2001US-0333347P.
XX XX (GEHO) GEN HOSPITAL CORP.
XX PA (FARB) BAYER AG.
XX PI Woolf C, D'urso D, Befort K, Costigan M;
XX XX
XX DR WPI; 2003-268312/26.
XX DR GENBANK; Q13183.
XX PT
XX PT New composition comprising two or more isolated polypeptides, useful for
XX preparing a medicament for treating pain in an animal.
XX
XX Claim 1; Page; 1017pp; English.
XX
XX The invention discloses a composition comprising two or more isolated rat
XX or human polynucleotides or a polynucleotide which represents a fragment,
XX derivative or allelic variation of the nucleic acid sequence. Also
XX claimed are a vector comprising the novel polynucleotide, a host cell
XX comprising the vector, a method for identifying a nucleotide sequence
XX which is differentially regulated in an animal subjected to pain and a
XX kit to perform the method, an array, a method for identifying an agent
XX that increases or decreases the expression of the polynucleotide sequence
XX that is differentially expressed in neuronal tissue of a first animal
XX subjected to pain, a method for identifying a compound which regulates
XX the expression of a polynucleotide sequence which is differentially
XX expressed in an animal subjected to pain, a method for identifying a
XX compound that regulates the activity of one or more of the
XX polynucleotides, a method for producing a pharmaceutical composition, a
XX method for identifying a compound or small molecule that regulates the
XX activity in an animal of one or more of the polypeptides given in the
XX specification, a method for identifying a compound useful in treating
XX pain and a pharmaceutical composition comprising the one or more
XX polypeptides or their antibodies. The polynucleotide or the compound that
XX modulates its activity is useful for preparing a medicament for treating
XX pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
XX injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
XX therapy). The sequence presented is a human protein (shown in Table 2 of
XX the specification) which is differentially expressed during pain. Note:
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic form directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
XX Sequence 592 AA;
XX

Query Match		78.2%; Score 2392; DB 7; Length 592;
Best Local Similarity		76.9%; Pred. No. 2e-232;
Matches		454; Conservative 63; Mismatches 67; Indels 6; Gaps 4;
QY	1	MATCWALWAYRYLVCLPIFLPLPLIVQKEAYCAYSIILMALLWCTEALPLAVTA 60
DB	1	MATCWALWAYRYLVCLPIFLPLPLIVQKEAYCAYSIILMALLWCTEALPLAVTA 60
QY	61	LPFVLPLMGIMDADEVCIYEKDTNIFVGLMVAIAVEHWNHKLRIALQVLLIIGVR 120
DB	61	LPFVLPLMGIMDADEVCIYEKDTNIFVGLMVAIAVEHWNHKLRIALQVLLIIGVR 120
QY	121	PALLLGFMVLVTAFLSNWISNTATTAMVPIGHAVLEQLQGSK--KDVGGNNNPTFELQ 178
DB	121	PAPLILGFMVLVTAFLSNWISNTATSAMVPIAHAVLDQLHSSQASSNVEGNNPTFELQ 180
QY	179	EECPQKEVTKLDNGQ--PV-SAPSEPTQKTOBHHRFSQGLSLCICYSASIGGIATLTGT 235
DB	181	EPSPQKEVTKLDNGQALPVTSSASSEGRAHLSQKHLHTQCMSLCVCYSASIGGIATLTGT 240
QY	236	TPNLVLQGVNSLFPQNGNVNFASFVGFAPFTMIILLALLWMLQVLFVGNFRKNFGF 295
DB	241	APNLVLQGVNSLFPQNGNVNFASFVGFAPFTMIILLALLWMLQVLFVGNFRKNFGF 300
QY	296	GEGERERKQAAFOVKTQYRLGLGPMGFAEKTVTVLVLLVLTFTREPGFFPGWGTVFA 355
DB	301	GERKQEQQAAYCVIQTEHRLGPMVFAEKAISILFVILVLLVLTFTREPGFFLGWGLAFP 360
QY	356	NEKGQWASDGTVAIFSLVFIPIPSKIPGLMDPKPKKAPPAIILTKTNDKMPWN 415
DB	361	NAGGESMVSDGTVAIFGIIMFIIPSKFGLTQDPENPGKAPGLGDLWKTQVQKMPWN 420
QY	416	IVLLGGGFALAKGSGSLSEWLGDKLTPLQHIPPSATAVILCLIAIFTECTSNVATT 475
DB	421	IVLLGGGFALAKGSGSLSEWLGDKLTPLQSVPAFAIILSLVATFTECTSNVATT 480
QY	476	TLFPLTILASMAQAICLHPLVYMLPCTILASLAPMLPVATPPNNAIVFSFGGLKVSMDARAG 535
DB	481	TIFLPLTILASMAQAICLHPLVYMLPCTILATSLAPMLPVATPPNNAIVFSFGDLKVLDMARAG 540
QY	536	FLANIIGVLAITISINWSIPFKLDTFPPSWAHSN-TSQCLLNPSTVTP 584
DB	541	FLANIIGVLIILAINSGIPLFSLSHSPWASQSNNTAQCLPSIANTTTP 590
RESULT 7		
ABM84184		
ID	ABM84184	standard; protein; 519 AA.
AC	ABM84184;	
XX		
DT	18-NOV-2004	(first entry)
XX		
DE		Human diagnostic and therapeutic pprotein SEQ ID NO:4433.
XX		
KW		gene therapy; human diagnostic and therapeutic polynucleotide; dithp.
XX		
OS		Homo sapiens.
XX		
PN	WO2004023973-A2.	
XX		
PD	25-MAR-2004.	
XX		
PF	12-SEP-2003; 2003WO-US028227.	
XX		
PR	12-SEP-2002; 2002US-0410259P.	
PR	12-SEP-2002; 2002US-0410260P.	
XX		
PA	(INCY-) INCYTE CORP.	
XX		
PI	Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;	
PI	Harthorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;	
PI	Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;	

PI	Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;	
PI	Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;	
PI	Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES;	
PI	Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;	
PI	Fatury S, Shi X, Suarez CU;	
XX		
DR	WPI; 2004-329368/30.	
DR	N-PSDB; ACN42836.	
XX		
PT	New diagnostic and therapeutic polynucleotides and polypeptides, useful	
PT	in diagnosing a condition, disease or disorder associated with human	
PT	molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or	
PT	in gene mapping.	
XX		
PS	Claim 27; Page; 190pp; English.	
XX		
CC	The invention relates to novel diagnostic and therapeutic polynucleotides	
CC	selected from one of the 2722 sequences defined in the specification. A	
CC	polynucleotide of the invention may have a use in gene therapy. The human	
CC	polynucleotide and therapeutic polynucleotides (dithp) or polypeptides may be	
CC	used to diagnose a particular condition, disease or disorder associated	
CC	with human molecules, e.g. cell proliferative disorders,	
CC	autoimmune/inflammatory disorder, developmental disorder, endocrine	
CC	disorder, neurological disorders, gastrointestinal disorders, or	
CC	infections caused by virus, bacteria, fungi or parasite. The dithp	
CC	molecules may also be used in genetic mapping, in identifying individuals	
CC	from minute biological samples, in detecting single nucleotide	
CC	polymorphisms, as molecular weight markers, and for somatic or germline	
CC	gene therapy. The present sequence represents a dithp protein of the	
CC	invention. Note: The sequence data for this patent is not represented in	
CC	the printed specification, but was obtained in electronic format directly	
CC	from WIPO at www.wipo.int/pct/en/sequences/listing.htm	
XX		
SQ	Sequence 519 AA;	
Query Match		67.8%; Score 2073.5; DB 8; Length 519;
Best Local Similarity		68.8%; Pred. No. 2.9e-200;
Matches		406; Conservative 48; Mismatches 57; Indels 79; Gaps 5;
QY	1	MATCHPALWAYRYLVCLPIFLPLIVQKEAYCAYSIILMALLWCTEALPLAVTA 60
DB	1	MATCHQALWAYRYLVCLPIFLPLIVQKEAYCAYSIILMALLWCTEALPLAVTA 60
QY	61	LPFVLPLMGIMDADEVCIYEKDTNIFVGLMVAIAVEHWNHKLRIALQVLLIIGVR 120
DB	61	LPFVLPLMGIMDADEVCIYEKDTNIFVGLMVAIAVEHWNHKLRIALQVLLIIGVR 120
QY	121	PALLLGFMVLVTAFLSNWISNTATTAMVPIGHAVLEQLQGSK--KDVGGNNNPTFELQ 178
DB	121	PAPLILGFMVLVTAFLSNWISNTATSAMVPIAHAVLDQLHSSQASSNVEGNNPTFELQ 180
QY	179	EECPQKEVTKLDNGQ--PV-SAPSEPTQKTOBHHRFSQGLSLCICYSASIGGIATLTGT 235
DB	181	EPSPQKEVTKLDNGQALPVTSSASSEGRAHLSQKHLHTQCMSLCVCYSASIGGIATLTGT 240
QY	236	TPNLVLQGVNSLFPQNGNVNFASFVGFAPFTMIILLALLWMLQVLFVGNFRKNFGF 295
DB	241	APNLVLQGVNSLFPQNGNVNFASFVGFAPFTMIILLALLWMLQVLFVGNFRKNFGF 293
QY	296	GEGERERKQAAFOVKTQYRLGLGPMGFAEKTVTVLVLLVLTFTREPGFFPGWGTVFA 355
DB	294	-----
QY	356	NEKGQWASDGTVAIFSLVFIPIPSKIPGLMDPKPKKAPPAIILTKTNDKMPWN 415
DB	294	-----MVSDGTVAIFGIIMFIIPSKFGLTQDPENPGKAPGLGDLWKTQVQKMPWN 347
QY	416	IVLLGGGFALAKGSGSLSEWLGDKLTPLQHIPPSATAVILCLIAIFTECTSNVATT 475
DB	348	IVLLGGGFALAKGSGSLSEWLGDKLTPLQSVPAFAIILSLVATFTECTSNVATT 407
QY	476	TLFPLTILASMAQAICLHPLVYMLPCTILASLAPMLPVATPPNNAIVFSFGGLKVSMDARAG 535

Db 408 TIFLILASMAQAICHLPLVYMLPCTLATSLAFMLPVATPPNAIVFSGDLKVLDMARAG 467

QY 536 FLNIIIGVLAITLSINSWSIPIFKLDTFPPSAHNS-TSQCLNPSNSTVP 584

Db 468 FLNIIIGVLAIALNSNGIPLFSLHSPSAQSNNTAQCLPSLANTTTP 517

RESULT 8

ID ADH22543 standard; protein; 519 AA.

AC ADH22543;

XX 11-MAR-2004 (first entry)

DE Human transporter & ion channel (TRICH) protein SeqID1.

KW human; transporters and ion channel; TRICH; cell proliferative;

KW arteriosclerosis; cancer; autoimmune/inflammatory; AIDS; asthma;

KW neurological; epilepsy; stroke; developmental; Cushing's syndrome;

KW hypothyroidism; infection; gene therapy; cytostatic; antiinflammatory;

KW immunosuppressive; antiaesthetic; anticonvulsant; nootropic;

KW neuroprotective; single nucleotide polymorphism; SNP.

XX Homo sapiens.

OS WO2003093444-A2.

PN 13-NOV-2003.

XX 02-MAY-2003; 2003WO-US014026.

XX 03-MAY-2002; 2002US-0377435P.

PR 03-MAY-2002; 2002US-0377444P.

PR 05-JUN-2002; 2002US-0386497P.

PR 11-JUN-2002; 2002US-0388180P.

XX (INCY-) INCYTE CORP.

XX Baughn MR, Becha SD, Bulloch SA, Chang H, Elliott VS;

PI Eméring BM, Griffin JA, Hafalata AJA, Ison CH, Jackson AA, Jiang X;

PI Jin P, Kable AE, Khare R, Lee SY, Lee S, Mason JM, Marquis JP;

PI Rankumar J, Richardson TW, Swarnakar A, Tran UK, Chawla NK;

PI Wilson AD;

XX WPI: 2004-022655/02.

DR N-PSDB; ADH22609.

XX New human transporters and ion channels (TRICH), useful for diagnosing,

PT treating and preventing diseases or conditions associated with the

PT aberrant TRICH expression e.g. cancer, AIDS, atherosclerosis, epilepsy,

PT or infections.

XX Claim 1; SEQ ID NO 41; 448pp; English.

XX This invention relates to novel isolated polynucleotides identified as

CC human transporters and ion channels (TRICH), and the encoded polypeptides

CC thereof. Specifically, it describes using these TRICH molecules, as well

CC as agonists, antagonists, antibodies, expression vectors and host cells,

CC in appropriate screening and toxicity assays to assess the effects of

CC exogenous compounds on TRICH expression. The present invention describes

CC TRICH compositions that are useful in the diagnosis, treatment and

CC prevention of various disorders such as cell proliferative (e.g.

CC arteriosclerosis, cancer), autoimmune/inflammatory (e.g. AIDS, asthma),

CC neurological (e.g. epilepsy, stroke) and developmental (e.g. Cushing's

CC syndrome hypothyroidism) and for infections. Accordingly, these TRICH

CC molecules can be used for gene therapy purposes and exhibit various

CC activities such as cytostatic, antiinflammatory, immunosuppressive,

CC antiaesthetic, anticonvulsant, nootropic and neuroprotective.

CC Furthermore, a microarray is useful in monitoring or measuring protein-

CC protein interactions, drug-target interactions and gene expression

CC profiles. This polypeptide sequence is a human TRICH protein of the

CC invention.

XX SQ Sequence 519 AA;

Query Match 67.5%; Score 2064.5; DB 8; Length 519;

Best Local Similarity 68.6%; Pred. No. 2.3e-199;

Matches 405; Conservative 48; Mismatches 58; Indels 79; Gaps 5;

QY 1 MATCPALWAYRFYLIIVLCPLIFLLPLPLIVQTKYACYSIILMALLMCTEALPLAVTA 60

Db 1 MATCQALWAYRSYLIIVFFVILLPLPLIPVPSKEYACAYAILMALLMCTEALPLAVTA 60

QY 61 LPFIVLPMLGIMDASEVCIIEYFKDTNLFVGGMLVAIAVEHNNLHKRIALQVLIIGVR 120

Db 61 LPFIVLPMLGIMDASEVCIIEYFKDTNLFVGGMLVAIAVEHNNLHKRIALQVLIIGVR 120

QY 121 PALLILGFMVLTAFLSMWISNTATTAMVPIGHAIVLEOLQSGK--KDYEGGNNNPFELQ 178

Db 121 PAPILGFMVLTAFLSMWISNTATTAMVPIGHAIVLEOLQSGK--KDYEGGNNNPFELQ 180

QY 179 EECQKEVTKLDNGQ--PV-SAPSEPRTKQTEHHRFSQGLSLCICYSASIGGIATLTGT 235

Db 181 EPSQKEVTKLDNGQALPVTSSASSEGRAHLSQKHLHLTQCHSLCVCYSASIGGIATLTGT 240

QY 236 TPNLVLOGQVNSLPQNGNVNPFASWFGFAFPTMIILLLLAWMLQVLFGLVNERKNEGF 295

Db 241 APNLVLOGQVNSLPQNGNVNPFASWFGFAFPTMIILLLLAWMLQVLFGLVNERKNEGF 293

QY 296 GEGEERKQAAQVQIKTQYRLLGPMSPFAEKTIVTVLVLLVVLVTRRPGPPFGWDTVPA 355

Db 294 ----- 293

QY 356 NEKGQSMASDGTVAIFISLVNFIIPSKIPLGMQDPKPKGLKAPPAIITWTKVNDKMPWN 415

Db 294 -----WVSDGTVAIFIGIIMFIIPSKFPGTQDPENPGKLPGLGLDWDKIVNQKMPWN 347

QY 416 IVILLGGFALAKGSEGLSEWLGDKLTPLQHIPPSATAVILCLLIAIFTCTSNVATT 475

Db 348 IVLLGGGYALAKGSEGLSEWLGDKLTPLQSVPAIPAIAIISLLVATFECTSNVATT 407

QY 476 TLFPLILASMAQAICHLPLVYMLPCTLATSLAFMLPVATPPNAIVFSGDLKVLDMARAG 535

Db 408 TIFLILASMAQAICHLPLVYMLPCTLATSLAFMLPVATPPNAIVFSGDLKVLDMARAG 467

QY 536 FLNIIIGVLAITLSINSWSIPIFKLDTFPPSAHNS-TSQCLNPSNSTVP 584

Db 468 FLNIIIGVLAIALNSNGIPLFSLHSPSAQSNNTAQCLPSLANTTTP 517

RESULT 9

ADG16995

ID ADG16995 standard; protein; 619 AA.

XX ADG16995;

XX 26-FEB-2004 (first entry)

DE African clawed frog transporter protein.

XX african clawed frog; transporter protein; pharmacogenomic analysis.

XX Xenopus laevis.

XX US2003186381-A1.

XX 02-OCT-2003.

XX 12-MAY-2003; 2003US-00435631.

XX 13-JUN-2000; 2000US-0211220P.

XX 05-DEC-2000; 2000US-00729094.

XX (APPL-) APPLERA CORP.

PI Chaturvedi K, Wei M, Ketchum KA, Difrancesco V, Beasley EM;
 XX WPI; 2004-032009/03.
 XX New isolated peptide for identifying its modulator, or an agent that
 PT binds to it, where the agent is used in a pharmaceutical composition for
 PT treatment of a disease or condition mediated by a human transporter
 PT protein.
 XX Disclosure; SEQ ID NO 4; 69pp; English.
 XX The invention relates to an isolated transporter protein. The invention
 CC is used for identifying its modulator, or an agent that binds to it,
 CC where agent is useful for a pharmaceutical composition for treatment of a
 CC disease or condition mediated by a human transporter protein. It is used
 CC as models for the development of human therapeutic targets; aid in the
 CC identification of therapeutic proteins; and serves as targets for the
 CC development of human therapeutic agents that modulate transporter
 CC activity, in cells and tissues that express the transporter. It is useful
 CC for development of commercially important products and services; and in
 CC pharmacogenomic analysis. The invention allows effective clinical design
 CC of treatment compounds and dosage regimens. The present sequence
 CC represents the amino acid sequence of an african clawed frog transporter
 CC protein.
 XX Sequence 619 AA;
 SQ
 Query Match 64.5%; Score 1971.5; DB 8; Length 619;
 Best Local Similarity 63.2%; Pred. No. 7.9e-190;
 Matches 386; Conservative 87; Mismatches 85; Indels 53; Gaps 10;
 QY 10 AVRFYLVLCPLPFLPLIVOTKAYCAYSIIIMALLWCTEALPLAVTALFPVLLFPL 69
 DB 10 ANRNYFIIFLPLPLPLVPLVFTKASCGFVIIVMALFWCTEALPLAVTALFPVLLFPM 69
 QY 70 MGIMDASEVCIEYFKDNLIFVGLMVAIAVEHNLHKLRIALQVLLIIGVRPALLLLGFM 129
 DB 70 MGIMDSTAVCSQYKLDNMLFVGLMVAIAVEHNLHKLRIALQVLLIIGVRPALLLLGFM 129
 QY 130 LVTAFLSMWISNTATTAMVPVIGHAVLEQLQSGK---KDVGGNN----- 171
 DB 130 VVTAFLSMWISNTATTAMVPVIGHAVLEQLQSGK---KDVGGNN----- 171
 QY 172 -----NPTFLOBEC-----POKEVTKLDNGQPVSAPEPRTQKT 206
 DB 190 SVNPSGKMAIAIDVTATENEGEIEQKSTKDEPSKQ-EKQSIGIVLIEPEDEKQTEK 248
 QY 207 Q--EHRFSQGLSLCICYSASIGGIATLTGTTNVLQGVNSLFPQNGVNVNFAWFGF 264
 DB 249 QKEKHLKICKGMSLCVCSASIGGIATLTGTTNVLQGVNSLFPQNGVNVNFAWFGF 308
 QY 265 AFTNMIILLLAWLQVLEFLGVNFRKNFGFGEGEE--ERKQAAFOVIKTOYELLGPMSP 322
 DB 309 AFTNMIILLLAWLQVLEFLGVNFRKNFGFGEGEE--ERKQAAFOVIKTOYELLGPMSP 368
 QY 323 AEKTVTVLVLLVLFVTRPFGFPGMGDTVFANEGQSMASDGTVAIFSLVMFIIIPSK 382
 DB 369 AEISLVLFLLVLLVLFVTRPFGFPGMGWA-TISFNKGGKEMVTATVAIFSLVMFIIIPSK 427
 QY 383 IPGL-MQDPKKPG---KLKAPPAITLTKTVNDKMPNIVILLGGFALAKGSGSLSEW 438
 DB 428 LPSFKYQDTPKGMKPKLKVPPALLDVKTNKMPNIVILLGGFALAKGSGSLSEW 487
 QY 439 LGKLPFLQHPSTATAVILICILIAITFECTSNVATTTTFLPLASNAQAICLHPVYML 498
 DB 488 LGKLPFLQHPSTATAVILICILIAITFECTSNVATTTTFLPLASNAQAICLHPVYML 547
 QY 499 PCTLASLAFMLPVATPPNAIVFSFGGLKVDSDMARAGFLNIIGVLAITLSINSWSPIF 558
 DB 548 PCTLASLAFMLPVATPPNAIVFSFGGLKVDSDMARAGFLNIIGVLAITLSINSWSPIF 607
 QY 559 KLDTPFSAWHS 569
 | | | | | | | | | |

DB 608 NLGTFPSWANA 618
 RESULT 10
 ADP64799
 ID ADP64799 standard; protein; 581 AA.
 XX AC ADP64799;
 XX DT 26-AUG-2004 (first entry)
 XX DE Zebrafish Na+-coupled citrate transporter protein.
 DE sodium-coupled citrate transporter; transmembrane citrate transporter;
 KW lifespans; weight reduction; weight gain prevention; blood cholesterol;
 KW triglyceride; low density lipoprotein; glucose; obesity;
 KW hyperlipidemia; hypercholesterolemia; INDI protein.
 XX Danio rerio.
 OS WO2004048925-A2.
 PN 10-JUN-2004.
 PD 20-NOV-2003; 2003WO-US037054.
 PF 22-NOV-2002; 2002US-0428469P.
 PR 01-APR-2003; 2003US-0459441P.
 XX (MEDI-) MEDICAL COLLEGE GEORGIA RES INST.
 PA (GANA/) GANAPATHY V.
 PA (INOU/) INOUE K.
 PA (FEIY/) FEI Y.
 PI Ganapathy V, Inoue K, Fei Y;
 XX WPI; 2004-460797/43.
 DR N-PSDB; ADP64798.
 XX New isolated polynucleotide encoding a Na+-coupled citrate transporter
 PT (NaCT) polypeptide, useful as a drug target for the treatment of obesity,
 PT hyperlipidemia, and hypercholesterolemia.
 XX Claim 42; SEQ ID NO 12; 186pp; English.
 CC The invention relates to novel Na+-coupled citrate transporter proteins
 CC and their encoding genes. Inhibitors of transmembrane citrate
 CC transporters are useful for extending the lifespan, reducing weight,
 CC preventing weight gain or lowering blood cholesterol, triglyceride, LDL
 CC or glucose levels in a subject. The NaCT polypeptide is useful as a drug
 CC target for the treatment of obesity, hyperlipidemia, and
 CC hypercholesterolemia. This sequence corresponds to the zebrafish Na+-
 CC coupled citrate transporter protein.
 XX Sequence 581 AA;
 SQ
 Query Match 59.8%; Score 1829; DB 8; Length 581;
 Best Local Similarity 60.4%; Pred. No. 1.9e-175;
 Matches 349; Conservative 92; Mismatches 113; Indels 24; Gaps 6;
 QY 8 LWAYRYLVLCPLPFLPLIVOTKAYCAYSIIIMALLWCTEALPLAVTALFPVLLF 67
 DB 9 VWKMKNTLILFCTPFLPLPLVIGSKEAGCAVYVLMVAIYVCTEVLPLAVTALLPAVL 68
 QY 68 PLWGMINDASEVCIEYFKDNLIFVGLMVAIAVEHNLHKLRIALQVLLIIGVRPALLLLG 127
 DB 69 PLRIMESQDVCQYKLDNMLFVGLMVAIAVEHNLHKLRIALQVLLIIGVRPALLMLG 128
 QY 128 FMLVTAFLSMWISNTATTAMVPVIGHAVLEQLQSGKDVGG---NNNPTFELQEECPQK 184
 DB 129 FMCVTAFLSMWISNTATTAMVPVIGHAVLEQLQSGKDVGG---NNNPTFELQEECPQK 188
 QY 185 EVTKLDNGQPVSAPEP--RTQKTQHHRPSQGLSLCICYSASIGGIATLTGTTNVLQ 242

Db 189 EKVL-NGDNFWSDEPESHSREAEERLKWSKGLTLCVYASIGTALTGTGPNVLUM 247
Qy 243 GQNSLFPQNGVNVFASFGFAFPMTIILLALLMLQLVFLGVNFRKNFGFGEGER 302
Db 248 GQMSQLFPNDPDIINFASFGFAFPNMIIMLTAWLWLQIVFLGINFKKTGCGTVKTEK 307
Qy 303 KQAAQVQIKTVYRLGPNMSFAKTVTVLVLVWLFTRPFGFFGWDVTVFANEKGQSM 362
Db 308 EIAANVNIKEHRSGLPMTFGSLVLFLLVWLFTRDPGFDVGDWATRRFNADK--EF 365
Qy 363 ASDGTVAIFISLWMTIPISKIPGLM-----QDPKPKGKLKAPPAIITWTVNDK 411
Db 366 VDATVAVFAALLFVFPKPRLCFWRTESDTPVQSSGP-----TPALLTWTQTKK 420
Qy 412 MPWNIVILGGGFALAKGSEQSLSEWLGDKLTPLQHIPPSATAVILCILIAIFTECTSN 471
Db 421 MPWSIILLGGGFALAKGSEISGLSKWLGQDQSPQSQIPPPWAIIVICLMIATFTECTSN 480
Qy 472 VATTTLFLPILASMAQAICLHPLYVMLPCTLAASLAFMLPVATPPNAIVFSFGGLKVSDM 531
Db 481 VATATLFLPILASMSQISQVNPPLYVMPCTLSASFAMLPVATPPNAIVFSYGLKVSMD 540
Qy 532 ARAGELLNIGVLAITLSINSISPIFKLDTFPSWAHS 569
Db 541 AKTGIVNIIIGLITSLAINSGRAIFSLDTFPSWANT 578

RESULT 11
ABU65064
ID ABU65064 standard; protein; 616 AA.
AC ABU65064;
DT 20-MAY-2003 (first entry)
DE Human NOVI4c protein.
KW NOVI; cytostatic; cardiant; antiarteriosclerotic; antiasthmatic; cancer;
KW hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;
KW human.
XX Homo sapiens.
OS WO200272757-A2.
PN 19-SEP-2002.
PD 08-MAR-2002; 2002WO-US006908.
PF 08-MAR-2001; 2001US-0274101P.
PR 08-MAR-2001; 2001US-0274194P.
PR 08-MAR-2001; 2001US-0274281P.
PR 09-MAR-2001; 2001US-0274322P.
PR 12-MAR-2001; 2001US-0274849P.
PR 13-MAR-2001; 2001US-0275578P.
PR 13-MAR-2001; 2001US-0275579P.
PR 13-MAR-2001; 2001US-0275601P.
PR 14-MAR-2001; 2001US-0276000P.
PR 16-MAR-2001; 2001US-0276776P.
PR 19-MAR-2001; 2001US-0276994P.
PR 20-MAR-2001; 2001US-0277239P.
PR 20-MAR-2001; 2001US-0277321P.
PR 20-MAR-2001; 2001US-0277327P.
PR 21-MAR-2001; 2001US-0277791P.
PR 22-MAR-2001; 2001US-0277833P.
PR 23-MAR-2001; 2001US-0278152P.
PR 26-MAR-2001; 2001US-0278894P.
PR 27-MAR-2001; 2001US-0278999P.
PR 27-MAR-2001; 2001US-0279036P.
PR 28-MAR-2001; 2001US-0279344P.
PR 30-MAR-2001; 2001US-0277338P.

PR 30-MAR-2001; 2001US-0279995P.
PR 30-MAR-2001; 2001US-0280233P.
PR 02-APR-2001; 2001US-0280802P.
PR 02-APR-2001; 2001US-0280822P.
PR 02-APR-2001; 2001US-0280900P.
PR 04-APR-2001; 2001US-0281194P.
PR 13-APR-2001; 2001US-0283675P.
PR 30-APR-2001; 2001US-0287424P.
PR 03-MAY-2001; 2001US-0288066P.
PR 03-MAY-2001; 2001US-0288342P.
PR 03-MAY-2001; 2001US-0288528P.
PR 15-MAY-2001; 2001US-0291190P.
PR 16-MAY-2001; 2001US-0291099P.
PR 16-MAY-2001; 2001US-0291240P.
PR 30-MAY-2001; 2001US-0294485P.
PR 31-MAY-2001; 2001US-0294889P.
PR 31-MAY-2001; 2001US-0294899P.
PR 18-JUN-2001; 2001US-0299027P.
PR 19-JUN-2001; 2001US-0299303P.
PR 19-JUN-2001; 2001US-0299310P.
PR 10-JUL-2001; 2001US-0304354P.
PR 31-JUL-2001; 2001US-0309198P.
PR 16-AUG-2001; 2001US-0312903P.
PR 10-SEP-2001; 2001US-0318462P.
PR 12-SEP-2001; 2001US-0318770P.
PR 27-SEP-2001; 2001US-0325430P.
PR 27-SEP-2001; 2001US-0325681P.
PR 18-OCT-2001; 2001US-0330380P.
PR 31-OCT-2001; 2001US-0335301P.
PR 14-NOV-2001; 2001US-0332172P.
PR 14-NOV-2001; 2001US-0332271P.
PR 14-NOV-2001; 2001US-0333184P.
PR 14-NOV-2001; 2001US-0333272P.
PR 21-NOV-2001; 2001US-0332094P.
PR 03-DEC-2001; 2001US-0337426P.
PR 03-DEC-2001; 2001US-0338092P.
PR 04-DEC-2001; 2001US-0337185P.
PR 03-JAN-2002; 2002US-0345705P.
PR 07-MAR-2002; 2002US-00092900.
XX (CURA-) CURAGEN CORP.
PA Padigar M, Spytek KA, Shenoy SG, Taupier RJ, Pena CEA, Li L;
XX Zehruseen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;
PI Patturajan M, Gangolli E, Vernet CAM, Guo X, Tchernev V;
PI Fernandes ER, Casman SJ, Malyankar UM, Gerlach V, Liu Y, Anderson D;
PI Spaderna SK, Catterton E, Burgess C, Leite M, Zhong H, Alsobrook JP;
PI Lepley DM, Rieger DK;
XX MPI; 2002-723332/78.
DR N-PSDB; ABX97031.
XX NOVI polypeptides and polynucleotides, useful for preventing or treating
PT a disorder associated with aberrant NOVI expression or activity e.g.,
PT cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
PT asthma.
XX Claim 1; Page 134; 1103pp; English.
PS This invention describes novel human NOVI polypeptides which have
XX cytostatic, cardiant, antiarteriosclerotic, antiasthmatic and hypotensive
CC activity. Pharmaceutical compositions comprising the NOVI proteins or
CC nucleic acid molecules or NOVI antibodies are useful for preventing or
CC treating a disorder associated with aberrant NOVI expression or activity
CC e.g. cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
CC asthma. The products of the invention can be used for gene therapy or in
CC a vaccine. ABU65041-ABU65218 represent the NOVI polypeptides encoded by
CC ABX97008-ABX97185
XX Sequence 616 AA;
SQ Query Match 53.2%; Score 1628; DB 5; Length 616;

Best Local Similarity 51.3%; Pred. No. 4.6e-155;		
Matches 314; Conservative 113; Mismatches 139; Indels 46; Gaps 7;		
Qy	1	MATCWPAALWAYRYLVLCPLPIPELLPLPIVQTKYACAYSIILMALLWCTEALPLAVTA 60
Db	1	MASALSYVSKFSFVILFVTPLLLPLVILMPAKVSCCAYVILMAIYWCTEVIPLEAVTS 60
Qy	61	LFPVLFPPLMGIMDASVCIEYFKDTNILFVGGLMVAIAVEHNLKRIALQVLLIIGVR 120
Db	61	LMPVLFPPLFQIILDSROVCVQYKDTNMLFGLIIVAVAVERNLKHRIALRILLWVGAK 120
Qy	121	PALLLGFMLVTAFLSMWISNTATTAMVPIGHVLEQLQSGKQVEGG-----NNN 172
Db	121	PARMLGFMGVGTALLSMWISNTATTAMVPIVEAILQMEATSAATEAGLEGQGTINNL 180
Qy	173	PTPE-----LQECPO-----KEVTKL-----DNGQPVSAPEP 201
Db	181	NALEDDTVKAVLGKCVAAISTYVKVEKLQINNMTPLKKLEKQEQDILGPIRPQDSA 240
Qy	202	RTOKTOGHRHFSOGLSLICYSASIGGIATLTGTPNLVLOGOVNSLFPONGNVNFAFW 261
Db	241	QCQEDQBRKRLKAMTLCICYAASIGGTALTGTGPNVLLGQWNELFPPDSKDLNFAFW 300
Qy	262	FGFAFTMIILLIALLMWLQVLGNFRKNFGFGEGERKQAQFQVIKTYRLILGPMS 321
Db	301	FATFAPNMLVMLLFAWLQFVVMFSFKSGWGLESKKNEKAALKVLQEEYVKLGPLS 360
Qy	322	FASKTVTVLVLLVWLTREPGFFPGWDTVFANEKGQSMASDGTVAIFISLVMFIIPS 381
Db	361	FAEINVLCIFLLVILWFSRDPGFMGPGLTVAWVEGETKSV-SDATVAIFVATLLFIIPS 419
Qy	382	KIP-----GLMQPKPKGLKAPPAITKTWVDKMPNIVILLGGFALAKSEOSGLS 436
Db	420	QKAPNFRSQTEGKSP-VLIAPPLLDWKVTQEKVPWGIIVLLGGFALAKSEASGLS 478
Qy	437	EWLGDKLTPLOHPPSATAVILCLLIAIFTCTSNVATTLFLPILASMAQAICLHPLYV 496
Db	479	VWVGKQMEPLHAVPPAAITLILSLVAVTCTSNVATTLFLPILASMSRSGILNPLYI 538
Qy	497	MLPCTLAASAFMLPVATPPNAIVFSGGLKVSMDARAGFLNLTIGVLAITLINSWSIP 556
Db	539	MLPCTLSASAFMLPVATPPNAIVFTYGHKLVADMKVTGVMNIIIGVFCVFLAVNTWGRA 598
Qy	557	IFKLDTPPSNAH 568
Db	599	IFDLDFPDWAN 610
RESULT 12		
ID	ABU65063	
ID	ABU65063 standard; protein; 616 AA.	
AC	ABU65063;	
DT	20-MAY-2003 (first entry)	
DE	Human NOV14b protein.	
KW	NOVX; cytostatic; cardiant; antiarteriosclerotic; antiasthmatic; cancer;	
KW	hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;	
KW	human.	
OS	Homo sapiens.	
PN	WO200272757-A2.	
PD	19-SEP-2002.	
PF	08-MAR-2002; 2002WO-US006908.	
PR	08-MAR-2001; 2001US-0274101P.	
PR	08-MAR-2001; 2001US-0274194P.	
PR	08-MAR-2001; 2001US-0274281P.	

PR	08-MAR-2001; 2001US-0274322P.	
PR	09-MAR-2001; 2001US-0274849P.	
PR	12-MAR-2001; 2001US-0275235P.	
PR	13-MAR-2001; 2001US-0275578P.	
PR	13-MAR-2001; 2001US-0275579P.	
PR	13-MAR-2001; 2001US-0275601P.	
PR	14-MAR-2001; 2001US-0276000P.	
PR	16-MAR-2001; 2001US-0276776P.	
PR	19-MAR-2001; 2001US-0276994P.	
PR	20-MAR-2001; 2001US-0277239P.	
PR	20-MAR-2001; 2001US-0277321P.	
PR	20-MAR-2001; 2001US-0277327P.	
PR	21-MAR-2001; 2001US-0277791P.	
PR	22-MAR-2001; 2001US-0277833P.	
PR	23-MAR-2001; 2001US-0278152P.	
PR	26-MAR-2001; 2001US-0278894P.	
PR	27-MAR-2001; 2001US-0278999P.	
PR	27-MAR-2001; 2001US-0279036P.	
PR	28-MAR-2001; 2001US-0279344P.	
PR	30-MAR-2001; 2001US-0277338P.	
PR	30-MAR-2001; 2001US-0279959P.	
PR	30-MAR-2001; 2001US-0280233P.	
PR	02-APR-2001; 2001US-0280802P.	
PR	02-APR-2001; 2001US-0280822P.	
PR	02-APR-2001; 2001US-0280900P.	
PR	04-APR-2001; 2001US-0281194P.	
PR	13-APR-2001; 2001US-0283675P.	
PR	30-APR-2001; 2001US-0287424P.	
PR	02-MAY-2001; 2001US-0288066P.	
PR	03-MAY-2001; 2001US-0288342P.	
PR	03-MAY-2001; 2001US-0288528P.	
PR	15-MAY-2001; 2001US-0291190P.	
PR	16-MAY-2001; 2001US-0291099P.	
PR	16-MAY-2001; 2001US-0291240P.	
PR	30-MAY-2001; 2001US-0294485P.	
PR	31-MAY-2001; 2001US-0294889P.	
PR	31-MAY-2001; 2001US-0294899P.	
PR	18-JUN-2001; 2001US-0299027P.	
PR	19-JUN-2001; 2001US-0299303P.	
PR	19-JUN-2001; 2001US-0299310P.	
PR	10-JUL-2001; 2001US-0304354P.	
PR	31-JUL-2001; 2001US-0309198P.	
PR	16-AUG-2001; 2001US-0312903P.	
PR	10-SEP-2001; 2001US-0318462P.	
PR	12-SEP-2001; 2001US-0318770P.	
PR	27-SEP-2001; 2001US-0325430P.	
PR	27-SEP-2001; 2001US-0325681P.	
PR	18-OCT-2001; 2001US-0330380P.	
PR	31-OCT-2001; 2001US-0335301P.	
PR	14-NOV-2001; 2001US-0332172P.	
PR	14-NOV-2001; 2001US-0332711P.	
PR	14-NOV-2001; 2001US-0332722P.	
PR	14-NOV-2001; 2001US-0333184P.	
PR	14-NOV-2001; 2001US-0333272P.	
PR	21-NOV-2001; 2001US-0332094P.	
PR	03-DEC-2001; 2001US-0337426P.	
PR	03-DEC-2001; 2001US-0338092P.	
PR	04-DEC-2001; 2001US-0337185P.	
PR	03-JAN-2002; 2002US-0345705P.	
PR	07-MAR-2002; 2002US-00092900.	
PA	(CURA-) CURAGEN CORP.	
XX	Padigaru M, Spytek KA, Shenoy SG, Taupier RJ, Pena CEA, Li L;	
XX	Zerhusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;	
PI	Patturajan M, Gangolli E, Vernet CAM, Guo X, Tchernev V;	
PI	Fernandes ER, Casman SJ, Malyankar UM, Gerlach V, Liu Y, Anderson D;	
PI	Spaderna SK, Catterton E, Burgess C, Leite M, Zhong H, Alsobrook JP;	
PI	Lepley DM, Rieger DK;	
XX	WFI; 2002-723332/78.	
DR	N-PSDB; ABX97030.	
XX		

PT NOVX polypeptides and polynucleotides, useful for preventing or treating
PT a disorder associated with aberrant NOVX expression or activity e.g.,
PT cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
PT asthma.
PS Claim 1; Page 134; 1103pp; English.
XX This invention describes novel human NOVX polypeptides which have
CC cytostatic, cardiant, antiarteriosclerotic, antiasthmatic and hypotensive
CC activity. Pharmaceutical compositions comprising the NOVX proteins or
CC nucleic acid molecules or NOVX antibodies are useful for preventing or
CC treating a disorder associated with aberrant NOVX expression or activity
CC e.g. cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
CC asthma. The products of the invention can be used for gene therapy or in
CC a vaccine. AB065041-AB065218 represent the NOVX polypeptides encoded by
XX ABX97008-ABX97185
XX Sequence 616 AA;
SQ
Query Match 53.2%; Score 1628; DB 5; Length 616;
Best Local Similarity 51.3%; Pred. No. 4.6e-155;
Matches 314; Conservative 113; Mismatches 139; Indels 46; Gaps 7;
QY 1 MATCHPALWAYRFLVILCLPIFLPLPLIVQTKAYCAYSIIIMALLWCTEALPLAVTA 60
Db 1 MASALSYSKPKSFVILFVTPLELLPLVILMPAKVSCCAVYIIMAIYWCTEVPLAVTS 60
QY 61 LPFIVLPLMGIMDASEVCIEYFKDTNLFVGLMVAIAVEHNLHKKRIALQVLLIIGVR 120
Db 61 LMPVLLPFLQLDSRQVCVQYKMDTNMLFLGLIVAVAVERNHKKRIALRTLLWVCAK 120
QY 121 PALLLGFMLVAFLSMWISNTATTAMVPIGHAVLEQLQGSKKDVEGG-----NNN 172
Db 121 PARMLGFMGVTTALLSMWISNTATTAMVPIVEAILQOMEATSAATEAGLEGQGTINN 180
QY 173 PTFE-----LOECPQ-----KEVTKL-----DNGQPVSAPEP 201
Db 181 NAEEDDTTKAVLGKCKVAILSTYVYKVEKQINNNMTPLKKLEKQOQDLGFGIRPQSA 240
QY 202 RTQKQEHRRFSQGLSLCICYASIGGIATLTGTPNLVLOGVNSLPPONGVNVNFASW 261
Db 241 QCQEQERKRLCKAMTLCICVYASIGGTATLTGTGNVVLQGMELPDSKDLNVNFASW 300
QY 262 GFAPFTMIILLLAWMLQVFLGNFRKNFGGEBEERKQAAFOVIKQYRLLGPM 321
Db 301 FAFAPFNMLVLLFAWMLQVYMFSSFKSGGCGLESKNEKAAALKVLEBYRKLGPLS 360
QY 322 FAEKTVTVLVLLVLTREPGFPFGDTPVANEKGSQASDGTVAIFISLWFIIPS 381
Db 361 FAEINVLICFFLLVILWFSRDFGFMGLTVAWVEGETKSV-SDATVAIFVATLLFVPS 419
QY 382 KIP-----GLMQDPKPKGLKAPPAILTKVTNDKMPNIVILGGFPALAKGSQSGLS 436
Db 420 QPKPFNFSQTEEGKSP-VLIAPPPLDWKVTQEKVPMGIVLLGGFPALAKGSQSGLS 478
QY 437 EMLGDKLTPLQHPPSATAVILCLLIAIFTECTSNVATTTFLPLTASMAQICLHPLYV 496
Db 479 VMGQMEPLHAPVPAATILLSLLVAVFTECTSNVATTTFLPLFASMSRISGLNPLYI 538
QY 497 MLPCTLASLAPMLPVATPPNVAIVSFGLKVKSDMARAGFLNIIGVLAITLSINSWIP 556
Db 539 MLPCTLSASFAPMLPVATPPNVAIVFTYHGLKVDAMVKTGVMNITIGVFCVFLAVNTWGA 598
QY 557 IPKLDTPPSWAH 568
Db 599 IPDLDDHFDWAN 610
RESULT 13
ADK51048
ID ADK51048 standard; protein; 616 AA.
XX
AC ADK51048;

XX 17-JUN-2004 (first entry)
XX Human NOV18A protein sequence SeqID68.
XX cytostatic; NOVX-agonist; NOVX-antagonist; vaccine; gene therapy; cancer;
KW chromosome mapping; human; NOV18A.
XX Homo sapiens.
XX WO2003083046-A2.
XX 09-OCT-2003.
XX 01-APR-2003; 2003WO-US010142.
XX 02-APR-2002; 2002US-00115479.
XX 05-APR-2002; 2002US-0370349P.
XX 08-APR-2002; 2002US-0370969P.
XX 12-APR-2002; 2002US-0372019P.
XX 22-APR-2002; 2002US-0374379P.
XX 30-MAY-2002; 2002US-0384543P.
XX 03-JUN-2002; 2002US-00160619.
XX 15-AUG-2002; 2002US-0403748P.
XX 04-NOV-2002; 2002US-00287226.
XX 31-MAR-2003; 2003US-00403161.
XX (CURA-) CURAGEN CORP.
XX Anderson DW, Bento P, Boldog FL, Burgess CE, Casman SJ, Furtak K;
PI Gorman L, Gould-Rothberg BE, Gunther E, Heyes MP, Li L, Spytek KA;
PI Stone DJ, Zhong M, Malyankar UM, Edinger SR, Patturajan M;
PI Rothenberg ME, Smithson G;
XX WPI; 2003-812539/76.
XX N-PSDB; ADK51047.
PT New NOVX polypeptide, useful for preparing a composition for treating or
PT preventing e.g. cancer or for chromosome mapping.
XX Claim 1; SEQ ID NO 68; 433pp; English.
CC This invention relates to novel isolated polypeptides and the DNA
CC sequences which encode them. The invention may be useful for the
CC development of compounds with a cytostatic activity (as NOVX-agonists or
CC antagonists) or vaccines. In addition, the disclosed sequences may be
CC useful for gene therapy. The polypeptide is useful for preparing a
CC composition for treating or preventing a pathological state in a mammal,
CC for example cancer or for chromosome mapping. The present sequence is
CC that of a human NOVX protein of the invention.
XX Sequence 616 AA;
SQ
Query Match 53.2%; Score 1628; DB 7; Length 616;
Best Local Similarity 51.3%; Pred. No. 4.6e-155;
Matches 314; Conservative 113; Mismatches 139; Indels 46; Gaps 7;
QY 1 MATCHPALWAYRFLVILCLPIFLPLPLIVQTKAYCAYSIIIMALLWCTEALPLAVTA 60
Db 1 MASALSYSKPKSFVILFVTPLELLPLVILMPAKVSCCAVYIIMAIYWCTEVPLAVTS 60
QY 61 LPFIVLPLMGIMDASEVCIEYFKDTNLFVGLMVAIAVEHNLHKKRIALQVLLIIGVR 120
Db 61 LMPVLLPFLQLDSRQVCVQYKMDTNMLFLGLIVAVAVERNHKKRIALRTLLWVCAK 120
QY 121 PALLLGFMLVAFLSMWISNTATTAMVPIGHAVLEQLQGSKKDVEGG-----NNN 172
Db 121 PARMLGFMGVTTALLSMWISNTATTAMVPIVEAILQOMEATSAATEAGLEGQGTINN 180
QY 173 PTFE-----LOECPQ-----KEVTKL-----DNGQPVSAPEP 201
Db 181 NAEEDDTTKAVLGKCKVAILSTYVYKVEKQINNNMTPLKKLEKQOQDLGFGIRPQSA 240

CC	mapping, tissue typing, preventive medicine, and pharmacogenomics. The
CC	polypeptides are also useful as vaccines. This sequence represents an
CC	example of the polypeptide of the invention.
XX	
QQ	Sequence 616 AA;
Query Match	53.2%; Score 1628; DB 8; Length 616;
Best Local Similarity	51.3%; Pred. No. 4.6e-155;
Matches 314; Conservative 113; Mismatches 139; Indels 46; Gaps 7;	
QY	1 MATCPALWAYFYLVLCPLPLPLVOTKEAYCAYSIILMALLWCTALPLAVTA 60
DB	1 MASALSYVSKFSFVLFTPLLLPLVILMPAKVSCCAVILMAYWCTEVIPLAVTS 60
QY	61 LPPVLPPLMGINDASEVCIEYFKDITNLFVGLGMVAIAVEHNLHKLIALQVLIIGVR 120
DB	61 LMPVLLFPLFQILDSRQVCQVMKDTNMLFGLGLIVAVAVERNLHKLIALRLTLVWGAK 120
QY	121 PALLLGLFMLVTAFILSMWISNTATTAMVPIGHAVLEQLQGSKKDVEGG-----NNN 172
DB	121 PARMLGFMGVTAALLSMWISNTATTAMVPIVEAILQOMEATSAAEAGLEGQGTINNL 180
QY	173 PTFE-----LQEECPQ-----KEVTKL-----DNGQPVSAPEP 201
DB	181 NAELEDTVAVLGKCAVIAISTYVKKVEKLQNNLMTPLKLEKQEQDGLGPGIRPQDSA 240
QY	202 RTQKTEHRRFSQGLSLCICYSASIGGIATLTGTPNLVLOQVNSLPQNGNVNFWFASW 261
DB	241 QCQEQERKLCCKMTLCLCYAASIGGTATLTGTGNVLLQOMNELFPDSKDLNFWFASW 300
QY	262 FGFAPFTMIILLAWLMQVLFGLVNFNRFNGFGEERKQAAFOVTKTYRLLGPMWS 321
DB	301 FAFAPFNMLVLLFAWLQFVYMFSSPKSKGCGLESKKNEKAAALKVLEBYRKLGLPS 360
QY	322 FAEKTVTVLVLLVLTFTREBGFPGMGDTVPANEKGOSMASDGTVAIFISLVMPFIIPS 381
DB	361 FAELNVLCFFLLVILMFSDPGFMPGLTVAWVEGETKSV-SDATVAIFVATLLFIIPS 419
QY	382 KIP-----GLMDQPKPKGLKAPPAILTWKTVNDKMPNIVILGGGFALAKGSQSGLS 436
DB	420 QKPKFNFRSQTEGKSP-VLIAPPPLLDWKVQEKVPWGVILVLLGGGFALAKGSASGLS 478
QY	437 EWLQKLTPLQHPISATAVILCLIAIPTECTSNVATTLPLPIASMAQAIICLHPLYV 496
DB	479 VMGKQMEPLHAVPAAITLILSLVAVFECTSNVATTLPLPIFASMSRISGLNPLYI 538
QY	497 MLPCTLAASLAEWLPVATPPNAIVESFGCLKVSDMARAGFLNIIGVLAITLSINSWSIP 556
DB	539 MLPCTLSASFAPMLPVATPPNAIVFTYGHKLVADMVKTGVMNIIIGVFCVFLAVNTWGRA 598
QY	557 IPKLDTPFSSWAH 568
DB	599 IFDLDFPDWAN 610
RESULT 15	
ADN61777	
ID	ADN61777 standard; protein; 616 AA.
XX	
AC	ADN61777;
XX	
DT	01-JUL-2004 (first entry)
DE	Human novel protein NOV14b.
XX	
KW	Human; NOVX; diabetes; obesity; infectious disease; anorexia;
KW	cancer-associated cachexia; cancer; neurodegenerative disorder;
KW	Alzheimer's disease; Parkinson's disease; immune disorder;
KW	haematopoietic disorder; dyslipidaemia; chronic disease.
OS	Homo sapiens.
XX	
PN	US2004043382-A1.

XX	
PD	
XX	04-MAR-2004.
PF	
XX	07-MAR-2002; 2002US-00092900.
XX	
PR	08-MAR-2001; 2001US-0274191P.
PR	08-MAR-2001; 2001US-0274194P.
PR	08-MAR-2001; 2001US-0274281P.
PR	08-MAR-2001; 2001US-0274322P.
PR	09-MAR-2001; 2001US-0274849P.
PR	12-MAR-2001; 2001US-0275233P.
PR	13-MAR-2001; 2001US-0275578P.
PR	13-MAR-2001; 2001US-0275579P.
PR	13-MAR-2001; 2001US-0275601P.
PR	14-MAR-2001; 2001US-0276000P.
PR	16-MAR-2001; 2001US-0276776P.
PR	19-MAR-2001; 2001US-0276994P.
PR	20-MAR-2001; 2001US-0277239P.
PR	20-MAR-2001; 2001US-0277321P.
PR	20-MAR-2001; 2001US-0277327P.
PR	20-MAR-2001; 2001US-0277338P.
PR	21-MAR-2001; 2001US-0277791P.
PR	22-MAR-2001; 2001US-0277833P.
PR	23-MAR-2001; 2001US-0278152P.
PR	26-MAR-2001; 2001US-0278894P.
PR	27-MAR-2001; 2001US-0278999P.
PR	27-MAR-2001; 2001US-0279036P.
PR	28-MAR-2001; 2001US-0279344P.
PR	30-MAR-2001; 2001US-0279995P.
PR	30-MAR-2001; 2001US-0280233P.
PR	02-APR-2001; 2001US-0280802P.
PR	02-APR-2001; 2001US-0280822P.
PR	02-APR-2001; 2001US-0280900P.
PR	04-APR-2001; 2001US-0281444P.
PR	13-APR-2001; 2001US-0283675P.
PR	30-APR-2001; 2001US-0287424P.
PR	02-MAY-2001; 2001US-0288066P.
PR	03-MAY-2001; 2001US-0288342P.
PR	15-MAY-2001; 2001US-0291190P.
PR	16-MAY-2001; 2001US-0291099P.
PR	30-MAY-2001; 2001US-0294485P.
PR	31-MAY-2001; 2001US-0294889P.
PR	31-MAY-2001; 2001US-0294899P.
PR	18-JUN-2001; 2001US-0299027P.
PR	19-JUN-2001; 2001US-0299303P.
PR	19-JUN-2001; 2001US-0299310P.
PR	10-JUL-2001; 2001US-0304354P.
PR	31-JUL-2001; 2001US-0309198P.
PR	16-AUG-2001; 2001US-0312903P.
PR	10-SEP-2001; 2001US-0318462P.
PR	12-SEP-2001; 2001US-0318770P.
PR	27-SEP-2001; 2001US-0325430P.
PR	27-SEP-2001; 2001US-0325681P.
PR	18-OCT-2001; 2001US-0330380P.
PR	31-OCT-2001; 2001US-0335301P.
PR	14-NOV-2001; 2001US-0332172P.
PR	14-NOV-2001; 2001US-0332271P.
PR	14-NOV-2001; 2001US-0332272P.
PR	14-NOV-2001; 2001US-0333184P.
PR	21-NOV-2001; 2001US-0333272P.
PR	03-DEC-2001; 2001US-0332094P.
PR	03-DEC-2001; 2001US-0337426P.
PR	03-DEC-2001; 2001US-0338092P.
PR	04-DEC-2001; 2001US-0337185P.
PR	03-JAN-2002; 2002US-0345705P.
XX	
PA	(PADI/) PADIGARU M.
PA	(SPYT/) SPYTEK K A.
PA	(SHEN/) SHENOY S G.
PA	(TAUP/) TAUPIER R J.
PA	(PENA/) PENA C E A.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2005, 08:36:57 ; Search time 22.5692 Seconds

(without alignments)
1941.542 Million cell updates/sec

Title: US-10-017-479A-4

Perfect score: 3056

Sequence: 1 MATCHPALWAYRFLYLCL.....HSNTSQCLLNPSNVTPOGL 587

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 segs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/1/1aa/5A-COMB.pep:*

2: /cgn2_6/ptodata/1/1aa/5B-COMB.pep:*

3: /cgn2_6/ptodata/1/1aa/6A-COMB.pep:*

4: /cgn2_6/ptodata/1/1aa/6B-COMB.pep:*

5: /cgn2_6/ptodata/1/1aa/PCTUS-COMB.pep:*

6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2392	78.2	599	4	US-09-949-016-9866
2	1325	43.3	601	4	US-09-949-016-9977
3	1259	41.2	627	4	US-09-556-916-8
4	1257	41.1	627	4	US-09-556-916-2
5	1256	41.1	627	4	US-09-949-016-6840
6	1255.5	41.1	626	4	US-09-556-916-20
7	1254	41.0	627	4	US-09-556-916-10
8	1253.5	41.0	626	4	US-09-556-916-14
9	1252	40.9	627	4	US-09-556-916-4
10	1250.5	40.9	626	4	US-09-556-916-22
11	1248.5	40.8	626	4	US-09-556-916-15
12	1097	35.9	581	4	US-09-556-916-12
13	1095	35.8	581	4	US-09-556-916-6
14	1093.5	35.8	580	4	US-09-556-916-24
15	1091.5	35.7	580	4	US-09-556-916-18
16	1023	33.5	561	4	US-09-949-016-8161
17	776	25.4	527	4	US-09-602-787A-516
18	570.5	18.7	368	4	US-09-270-767-42029
19	522	17.1	470	4	US-09-543-681A-5952
20	493	16.1	233	4	US-09-489-847-176
21	475	15.5	335	4	US-09-602-787A-518
22	407.5	13.3	524	3	US-09-134-001C-5457
23	314.5	10.3	230	4	US-09-270-767-43713
24	307.5	10.1	180	4	US-09-270-767-42669
25	281.5	9.2	194	4	US-09-270-767-59102
26	279	9.1	421	4	US-09-248-796A-20749
27	250	8.2	547	4	US-09-489-039A-13843

28	247	8.1	132	4	US-09-270-767-57987	Sequence 57987, A
29	233	7.6	169	4	US-09-270-767-57286	Sequence 57286, A
30	233	7.6	596	4	US-09-902-540-13547	Sequence 13547, A
31	229	7.5	548	4	US-09-902-540-11870	Sequence 11870, A
32	228	7.5	624	4	US-09-543-681A-4343	Sequence 4343, A
33	212	6.9	614	4	US-09-489-039A-12605	Sequence 12605, A
34	210.5	6.9	493	4	US-09-540-236-2120	Sequence 2120, A
35	208	6.8	302	4	US-09-902-540-10445	Sequence 10445, A
36	205.5	6.7	694	4	US-09-252-991A-22637	Sequence 22637, A
37	196	6.4	494	4	US-09-543-681A-7033	Sequence 7033, A
38	169	5.5	478	3	US-09-134-001C-4637	Sequence 4637, A
39	155.5	5.1	449	4	US-09-328-352-7512	Sequence 7512, A
40	154	5.0	470	4	US-09-438-185A-209	Sequence 209, A
41	148	4.8	443	4	US-09-602-787A-532	Sequence 532, A
42	138.5	4.5	211	4	US-09-198-452A-1166	Sequence 1166, A
43	137.5	4.5	597	4	US-09-489-039A-14112	Sequence 14112, A
44	133	4.3	281	4	US-09-710-279-3284	Sequence 3284, A
45	128	4.2	524	4	US-09-198-452A-369	Sequence 369, A
46	128	4.2	524	4	US-09-438-185A-353	Sequence 353, A
47	125.5	4.1	430	3	US-09-134-001C-2981	Sequence 2981, A
48	125	4.1	515	4	US-09-869-433-2	Sequence 2, Appli
49	124	4.1	619	4	US-09-540-236-2377	Sequence 2377, A
50	120.5	3.9	437	4	US-09-543-681A-6984	Sequence 6984, A
51	117.5	3.8	398	4	US-09-489-039A-13159	Sequence 13159, A
52	117	3.8	603	3	US-09-097-889-23	Sequence 23, Appli
53	117	3.8	603	4	US-09-098-079-23	Sequence 23, Appli
54	116.5	3.8	526	4	US-09-543-681A-8223	Sequence 8223, A
55	116	3.8	518	4	US-09-252-991A-30478	Sequence 30478, A
56	115.5	3.8	441	4	US-09-489-039A-10612	Sequence 10612, A
57	115.5	3.8	502	4	US-09-252-991A-23327	Sequence 2327, A
58	113	3.7	429	4	US-08-311-731A-287	Sequence 287, A
59	113	3.7	363	4	US-09-270-767-42613	Sequence 42613, A
60	111	3.6	482	4	US-09-328-352-7784	Sequence 7784, A
61	108.5	3.5	616	4	US-09-489-039A-12599	Sequence 12599, A
62	108	3.5	167	4	US-09-248-796A-20740	Sequence 20740, A
63	108	3.5	319	4	US-09-270-767-57927	Sequence 57927, A
64	105.5	3.4	485	4	US-09-134-000C-6031	Sequence 6031, A
65	105.5	3.4	533	1	US-08-294-872-2	Sequence 2, Appli
66	105.5	3.4	533	5	PCT-US95-09823-2	Sequence 2, Appli
67	105	3.4	907	3	US-08-863-102-2	Sequence 2, Appli
68	105	3.4	996	4	US-09-252-991A-27018	Sequence 27018, A
69	103.5	3.4	345	4	US-09-902-540-9836	Sequence 9836, A
70	103.5	3.4	529	4	US-09-489-039A-8824	Sequence 8824, A
71	103.5	3.4	549	4	US-09-489-039A-13779	Sequence 13779, A
72	103	3.4	459	4	US-09-489-039A-10630	Sequence 10630, A
73	102.5	3.4	729	4	US-09-107-532A-3970	Sequence 3970, A
74	102	3.3	116	4	US-09-602-787A-520	Sequence 520, A
75	102	3.3	569	2	US-08-750-723A-2	Sequence 2, Appli
76	102	3.3	569	3	US-09-191-275-2	Sequence 2, Appli
77	101.5	3.3	470	4	US-09-328-352-8240	Sequence 8240, A
78	101.5	3.3	638	4	US-09-252-991A-18036	Sequence 18036, A
79	101	3.3	509	4	US-09-252-991A-23341	Sequence 23341, A
80	101	3.3	614	4	US-09-543-681A-4330	Sequence 4330, A
81	100.5	3.3	436	4	US-09-583-110-4729	Sequence 4729, A
82	100.5	3.3	440	4	US-09-107-433-5201	Sequence 5201, A
83	100.5	3.3	463	4	US-09-252-991A-25488	Sequence 25488, A
84	99.5	3.3	456	3	US-09-134-001C-3771	Sequence 3771, A
85	99	3.2	898	3	US-08-863-102-1	Sequence 1, Appli
86	99	3.2	898	3	US-08-863-102-4	Sequence 4, Appli
87	98.5	3.2	492	4	US-09-489-039A-8343	Sequence 8343, A
88	98.5	3.2	652	3	US-08-987-123-2	Sequence 2, Appli
89	97.5	3.2	402	4	US-09-489-039A-10525	Sequence 10525, A
90	97.5	3.2	413	4	US-09-543-681A-5608	Sequence 5608, A
91	97.5	3.2	552	4	US-09-248-796A-20151	Sequence 20151, A
92	97	3.2	647	4	US-09-489-039A-11893	Sequence 11893, A
93	97	3.2	641	4	US-09-543-681A-4453	Sequence 4453, A
94	96	3.1	477	4	US-09-252-991A-29825	Sequence 29825, A
95	96	3.1	514	4	US-09-252-991A-25845	Sequence 25845, A
96	95.5	3.1	391	4	US-09-710-279-1236	Sequence 1236, A
97	95.5	3.1	422	3	US-09-134-001C-4048	Sequence 4048, A
98	95.5	3.1	422	4	US-09-809-665A-73	Sequence 73, Appli
99	95.5	3.1	448	4	US-09-134-000C-3794	Sequence 3794, A
100	95.5	3.1	482	4	US-09-328-352-6084	Sequence 6084, A

ALIGNMENTS

RESULT 1
US-09-949-016-9866
; Sequence 9866, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9866
; LENGTH: 599
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-9866

Query Match 78.2%; Score 2392; DB 4; Length 599;
Best Local Similarity 76.9%; Pred. No. 6e-235;
Matches 454; Conservative 63; Mismatches 67; Indels 6; Gaps 4;

QY 1 MATCPALWAYRFLIIVLCPIPLPIVQTKAYCAYSIIIMALLMCTEALPLAVTA 60
DB 8 MATCQALWAYRSYLVFFVPIILLPLIPVPSKEAYCAVAIIMALFWCTEALPLAVTA 67

QY 61 LPFIVLPIMGIMDASEVCIEYFKDTNLLFVGGLMVAIAVEHNLKRIALQVLLIIGVR 120
DB 68 LPFLLPFPMGIVDASEVAYEYKDSNLLFPFGLLVVAIAVEHNLKRIALRVLLIIGVR 127

QY 121 PALLLGLMVLTAFLSMWISNTATTAMVPIGHAVLEQSGSK--KDVEGNNNPTELQ 178
DB 128 PAPILGLMVLTAFLSMWISNTATSMVPFAHVLQDLSQSSQSSNVEGNNPTELQ 187

QY 179 EECQKEVTKLDNGQ--PV-SAPSEPRQTKQEHRRFSQGLSICICYSASIGGIATLTGT 235
DB 188 EPSQKEVTKLDNGQALPVTSSASSEGRAHLSQKHLHTQCMSLCVCYSASIGGIATLTGT 247

QY 236 TPNVLQGVNSLFPQNGVNVNFAWFGFAPPTMIILLLAWLWLVFLGVNFRKNFGF 295
DB 248 APNLVLOGVNSLFPQNGVNVNFAWFGFAPPTMIILLLAWLWLVFLGVNFRKNFGI 307

QY 296 GEGEERKQAAFOVKTQYRLLGPMPSFAEKTVTVLVLLVLMFTREBPGFPFGMGDTVEA 355
DB 308 GEKMQEQQAAYCVIQTREHLLGPMTPFAKSAISILFVILVLLMFTREBPGFPGWNLAFP 367

QY 356 NEKGQSMASDGTVAIFSLVMFIIPSKIPGLMQDPKPKGKUKAPAILTKTVNDKMPWN 415
DB 368 NAKGESMVSDDGTVAIFIGIIFIIIPSKPPGLTQDPENPGKLPKPLGLLDMDKTVNQKMPWN 427

QY 416 IVILLGGFALAKSGEQSLSEWLGDKLTPLQHPPSATAVILCLLTAIFTECTSNVATT 475
DB 428 IVLLGGYALAKSGERSGLSEWLGDKLTPLQHPPSATAVILCLLTAIFTECTSNVATT 487

QY 476 TLFPLASMAQAICLHPYVYMLPCTLAASLAFMLPVATPPNAIVFSGGLKVSDMARAG 535
DB 488 TIFPLASMAQAICLHPYVYMLPCTLATSLAFMLPVATPPNAIVFSGGLKVLDMARAG 547

QY 536 FLNLTIGVLATLTSINWSIPIFKLDTTPPSWAHNS--TSQCLLNPSNSTVP 584
DB 548 FLNLTIGVLATLTSINWSIPIFKLDTTPPSWAHNS--TSQCLLNPSNSTVP 597

RESULT 2
US-09-949-016-9977
; Sequence 9977, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9977
; LENGTH: 601
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-9977

Query Match 43.3%; Score 1325; DB 4; Length 601;
Best Local Similarity 44.4%; Pred. No. 4.6e-126;
Matches 267; Conservative 125; Mismatches 167; Indels 42; Gaps 11;

QY 11 YRFYLIVLCPIPLPIVQTKAYCAYSIIIMALLMCTEALPLAVTALPVLPLM 70
DB 16 YRFLFVFTVLVLLPLPIVLTKEACAYTLFVATFWLTEALPLSVTALLPSLMLPMF 75

QY 71 GIMDASEVCIEYFKDTNLLFVGGLMVAIAVEHNLKRIALQVLLIIGVRPALLLGFML 130
DB 76 GIMPSKVASAYFKDPHLLIGVICLATSEIKWNLKRIALQVMMVGVNPNALTLGMS 135

QY 131 VTAFSLMWISNTATTAMVPIGHAVLEQSGSKDVE-----GNNNPTELOQEC-- 181
DB 136 STAFSLMWISNTATTAMVPFAHVLQDLSQSSQSSNVEGNNPTELQ 195

QY 182 ----PQKEVTK-----LDNGQ-----PVSAPSEPRQTKQEHRRFSQGLSICICYSAS 225
DB 196 HEINERKENTKPVPGVYNDTGKISSKVELEKNSGMETKYRTKKGHVTRKLTCLCIAYSST 255

QY 226 IGGIATLTGTPNVLQGVNSLFPQNGVNVNFAWFGFAPPTMIILLLAWLWLVFL 285
DB 256 IGGLTITGTSTNLIPAEYFNTRY-P-DCRCLNFGSWFTSFPAALILLLWLVQWLF 314

QY 286 GVNFRKNFGEGEERKQAAFOVKTQYRLLGPMPSFAEKTVTVLVLLVLMFTREBPGF 345
DB 315 GFNFKEFKCGKTKTVQKACAEVQKQGLPIRYQEIIVTLVLIIMALLWFSRDPGF 374

QY 346 FPGMGDTVFANEXQSGMASDGTVAIFSLVMFIIPSKIPGLMODPKPKKKA--PPAIL 403
DB 375 VPCW-SALFSEYFG--FATDSTVALLIGLIFLIPAK---TLTKTTGTGEIVAFDYSPLI 428

QY 404 TWKTVNDKMPWNIVILLGGFALAKSGEQSLSEWLGDKLTPLQHPPSATAVILCLLIA 463
DB 429 TWKEFQSFMPWDAILVGGGFALADGCESSGKWNKLSPLGSLPAWLIILISSLMVT 488

QY 464 IFTECTSNVATTTLFPLASMAQAICLHPYVYMLPCTLAASLAFMLPVATPPNAIVFSP 523
DB 489 SLTEVASNPATITLFPILSPLAEAHVNPYIILIPSTLCTSFALLPVANPENAI VFSY 548

QY 524 GGLKVSMDARAGFLNLTIGVLATLTSINWSIPIFKLDTTPPSWAHNS--TSQCLLNPSNSTVP 583
DB 549 GHLKVIDMWKAGLVNIGVAVVWMLGICTIVPMFDLYTYPWSAPA-----MSNETM 600

QY 584 P 584

Db 601 P 601

RESULT 3

US-09-556-916-8
; Sequence 8, Application US/09556916
; Patent No. 6548271

GENERAL INFORMATION:

; APPLICANT: Turner, Alex
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Sands, Arthur T.

; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins

; FILE REFERENCE: 8535-041-999

; CURRENT APPLICATION NUMBER: US/09/556,916

; CURRENT FILING DATE: 2000-04-21

; NUMBER OF SEQ ID NOS: 32

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 8

; LENGTH: 627

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-556-916-8

Query Match

Best Local Similarity 41.2%; Score 1259; DB 4; Length 627;

Matches 253; Conservative 125; Mismatches 172; Indels 64; Gaps 13;

QY 12 RYLVLCPLPFLPLPLIVOTKEAYCAYSIIIMALLMCTEALPLAVTALPFIIVPLPLMG 71

Db 11 RKLWVCPVLLPLPLPHPSSEASCAYLVITVAVYVSEAVPLGAAALVPAFLYPPFG 70

QY 72 IMDASEVIEYFKDNTILFVGLMVAIAVEHWNHKLRIALQVLLIIGVRPALLLGFMLV 131

Db 71 VLSNEVAAYFKNTLLVGVICVAAVEKNLHKLRIALRWLMAGAKPGMLLCFMC 130

QY 132 TAFLSMISNTATTAMVPIGHAVLEQLQGSK-KDVEGNNN-----PTF 175

Db 131 TTLLSMLSNTSTTAMVPIVEAVLQELVSADEQLVAGNSNTEBAEPISLDVKNQPSL 190

QY 176 EL---QEECPQKEVKLD-----NQP-----VSAPSEPR 202

Db 191 ELIFVNEDRSNADLTLLMHNENLVGPSITNPITKANOHQKKHPSOEKPOVLTPS-PR 249

QY 203 TQKT----QEH--RFSQGLSLCICYSASIGGIATLTGTPNLVLQGVNSLFPQNGV 256

Db 250 KQKLRKYRSHDDQMICKLSLSISYSATIGLTTIIGTSLIFLEHFNQYPA-AEVV 308

QY 257 NFASWFGFAPFTMIILLALLWLQVLFGLVFNFRKNFGEGEE-ERKQAAFQVKTQYR 315

Db 309 NFGTWFLSFPISLIMLVVFWMHVLFGLCNFKETCSLSKKKTKRQLSEKRIQEBYE 368

QY 316 LIGPMSFAEKTVTVLVLLVVLVFWTRPFGPFGDGVFANEKGOSMADGTVAIFISLV 375

Db 369 KLGDISYEPMTVGFPIILMTVLVFWTRPFGVPGWDS--FFEKKG--YRTDATVSFVLGFL 424

QY 376 MFIIPSKIP--GLMOPDKPKGLKAPPAITLTKVTNDKMPNIVILLGGGALAGSEOS 433

Db 425 LFLIPAKKPCFKKNDGENQHSLSGTSEIITWKDFQKTPMPEIIVLVGGVALAGSGSKS 484

QY 434 GLSEWLGDKLPLOHIPPASATAVILCLLIAIIFTECTSNVATTTLPLPLASMAQAICLHP 493

Db 485 GLSTWIGNQMLSLSPWAVTLACILVSVTEFVSNPATITIFPLILCSLSETLHPN 544

QY 494 LYVMLPCTLAASLAPLVPATPPNAIVSFGLKVDMDARAGFLNITIGVLAITLSINSW 553

Db 545 LYTLPVTWCISFAVMLPVGNPPNAIVFSYGHQIKDMVKAGLVNIGLVVMVAINTW 604

QY 554 SIPIFKLDTFPSWA 567

Db 605 GVSFLHLDITYPAWA 618

RESULT 5

RESULT 4

US-09-556-916-2

; Sequence 2, Application US/09556916

; Patent No. 6548271

GENERAL INFORMATION:

; APPLICANT: Turner, Alex
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Sands, Arthur T.

; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins

; FILE REFERENCE: 8535-041-999

; CURRENT APPLICATION NUMBER: US/09/556,916

; CURRENT FILING DATE: 2000-04-21

; NUMBER OF SEQ ID NOS: 32

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 2

; LENGTH: 627

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-556-916-2

Query Match

Best Local Similarity 41.1%; Score 1257; DB 4; Length 627;

Matches 253; Conservative 124; Mismatches 173; Indels 64; Gaps 13;

QY 12 RYLVLCPLPFLPLPLIVOTKEAYCAYSIIIMALLMCTEALPLAVTALPFIIVPLPLMG 71

Db 11 RKLWVCPVLLPLPLPHPSSEASCAYLVITVAVYVSEAVPLGAAALVPAFLYPPFG 70

QY 72 IMDASEVIEYFKDNTILFVGLMVAIAVEHWNHKLRIALQVLLIIGVRPALLLGFMLV 131

Db 71 VLSNEVAAYFKNTLLVGVICVAAVEKNLHKLRIALRWLMAGAKPGMLLCFMC 130

QY 132 TAFLSMISNTATTAMVPIGHAVLEQLQGSK-KDVEGNNN-----PTF 175

Db 131 TTLLSMLSNTSTTAMVPIVEAVLQELVSADEQLVAGNSNTEBAEPISLDVKNQPSL 190

QY 176 EL---QEECPQKEVKLD-----NQP-----VSAPSEPR 202

Db 191 ELIFVNEDRSNADLTLLMHNENLVGPSITNPITKANOHQKKHPSOEKPOVLTPS-PR 249

QY 203 TQKT----QEH--RFSQGLSLCICYSASIGGIATLTGTPNLVLQGVNSLFPQNGV 256

Db 250 KQKLRKYRSHDDQMICKLSLSISYSATIGLTTIIGTSLIFLEHFNQYPA-AEVV 308

QY 257 NFASWFGFAPFTMIILLALLWLQVLFGLVFNFRKNFGEGEE-ERKQAAFQVKTQYR 315

Db 309 NFGTWFLSFPISLIMLVVFWMHVLFGLCNFKETCSLSKKKTKRQLSEKRIQEBYE 368

QY 316 LIGPMSFAEKTVTVLVLLVVLVFWTRPFGPFGDGVFANEKGOSMADGTVAIFISLV 375

Db 369 KLGDISYEPMTVGFPIILMTVLVFWTRPFGVPGWDS--FFEKKG--YRTDATVSFVLGFL 424

QY 376 MFIIPSKIP--GLMOPDKPKGLKAPPAITLTKVTNDKMPNIVILLGGGALAGSEOS 433

Db 425 LFLIPAKKPCFKKNDGENQHSLSGTSEIITWKDFQKTPMPEIIVLVGGVALAGSGSKS 484

QY 434 GLSEWLGDKLPLOHIPPASATAVILCLLIAIIFTECTSNVATTTLPLPLASMAQAICLHP 493

Db 485 GLSTWIGNQMLSLSPWAVTLACILVSVTEFVSNPATITIFPLILCSLSETLHPN 544

QY 494 LYVMLPCTLAASLAPLVPATPPNAIVSFGLKVDMDARAGFLNITIGVLAITLSINSW 553

Db 545 LYTLPVTWCISFAVMLPVGNPPNAIVFSYGHQIKDMVKAGLVNIGLVVMVAINTW 604

QY 554 SIPIFKLDTFPSWA 567

Db 605 GVSFLHLDITYPAWA 618

```
US-09-949-016-6840
; Sequence 6840, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CU001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6840
; LENGTH: 627
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6840

Query Match      41.1%; Score 1256; DB 4; Length 627;
Best Local Similarity 41.2%; Pred. No. 5.4e-119;
Matches 253; Conservative 124; Mismatches 173; Indels 64; Gaps 13;

QY 12 RFYLIVLCPIELLPLIVQTKKAYCAYSIIIMALLWCTEALPLAVTALFPVLPPLMG 71
DB 11 RKLVLVVCVPLLLPLPLVHPSEASCAYLVITAVYVWSEAVPLGAAALVPAFLYPPFG 70
QY 72 IMDASEVCIFYKDNITLFGVGLMVAIAVEHNLHKKRIALQVLLIIGVRPALLLGFMLV 131
DB 71 VLRSNEVAEYFKNTLLLVGVICVAAAEKKNLHKKRIALRMVLMAGKPGMLLCCFMC 130
QY 132 TAFLSMWISNTATTAMVPIGHAVLEQLQSGK-KOVEGGNN-----PTF 175
DB 131 TTLLSMWLSNTSTTAMVPIVEAVLQELVSAEDQLVAGNSNTEEAAPISLDVKNQPSL 190
QY 176 EL---QECPQKEVTKLDNGQPVSA-----NGQP-----VSAPSEPR 202
DB 191 ELIFVNEEDRSNADTLTMHNENLNGVPSITNPIKTANQHQKQHPQSEKPVLTSP-PR 249
QY 203 TQKT-----QEHH--RFSQGLSLCICYSASIGGIATLTGTPNLVLQGVNSLPQNGNV 256
DB 250 KOKLRKYRSHDDQMIKCLSLISYSATIGGLTTIIGTSTSLIFLHFNNQYPA-AEVV 308
QY 257 NFASWFGFAPPTMIILLLAWLQVLFGLVFNFRKNFGFGEGBE-ERKQAAFOVIKTOYR 315
DB 309 NFGTWFLSPFISLIMLVSVFWMHVLFLGCNFKETCSLKKKTKREQLSEKRIQEEYE 368
QY 316 LLGPMSPAECTVTVLVLLVLMVTFREPFGPPGMDTVFANEKQSMASDGTVAIFISLV 375
DB 369 KIGDISYEMVTGFFILMTVLMVTFREPFGVPGWDS--FFEKKG--YRTDATVSVFLGFL 424
QY 376 MFIIPSKIP--GLMQDPKPKGLKAPPAILTWTKTVDNKPWNIVILLGGGFALAKGSEQS 433
DB 425 LFLIPAKKPCPKGKNDGENQBSLGTESIITKDPQKTMPEIIVILVGGGYALASGSKS 484
QY 434 GLSWMGDKLTPLQHPISPSATAVILCLLIAIIFTECTSNVATTTLFLPILASMAQICLHP 493
DB 485 GLSTWIGNQMLSSLPWPVATVILACILSVISVTEFVNPNATITIFLILCSLSETMHPN 544
QY 494 LYVNLPTCLTAASLAPLVPATPPNAIVFSFGGLKVSDMARAGFLNLIIGVLAITLSINSW 553
DB 545 LYTLPVTMCISFAVMLPVGNPPNAIVFSYGHQCIKDMVKAGLVNIGLVIMVAINTW 604
QY 554 SIPIKLDTPPSWA 567
DB 605 GVSFLHLDTPAWA 618
```

```
RESULT 6
US-09-556-916-20
; Sequence 20, Application US/09556916
; Patent No. 6548271
; GENERAL INFORMATION:
; APPLICANT: Turner, Alex
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: NO. 6548271el Human Transporter Proteins
; FILE REFERENCE: 8535-041-999
; CURRENT APPLICATION NUMBER: US/09/556,916
; CURRENT FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 20
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-556-916-20

Query Match      41.1%; Score 1255.5; DB 4; Length 626;
Best Local Similarity 40.8%; Pred. No. 6e-119;
Matches 250; Conservative 123; Mismatches 178; Indels 61; Gaps 11;

QY 12 RFYLIVLCPIELLPLIVQTKKAYCAYSIIIMALLWCTEALPLAVTALFPVLPPLMG 71
DB 11 RKLVLVVCVPLLLPLPLVHPSEASCAYLVITAVYVWSEAVPLGAAALVPAFLYPPFG 70
QY 72 IMDASEVCIFYKDNITLFGVGLMVAIAVEHNLHKKRIALQVLLIIGVRPALLLGFMLV 131
DB 71 VLRSNEVAEYFKNTLLLVGVICVAAAEKKNLHKKRIALRMVLMAGKPGMLLCCFMC 130
QY 132 TAFLSMWISNTATTAMVPIGHAVLEQLQSGK-KOVEGGNN-----PTF 175
DB 131 TTLLSMWLSNTSTTAMVPIVEAVLQELVSAEDQLVAGNSNTEEAAPISLDVKNQPSL 190
QY 176 EL---QECPQKEVTKLDNGQPVSA-----EPRTQ 204
DB 191 ELIFVNEEDRSNADTLTMHNENLNGVPSITNPIKTANQHQKQHPQSEKPVLTSPRQ 250
QY 205 KT-----QEHH--RFSQGLSLCICYSASIGGIATLTGTPNLVLQGVNSLPQNGNVN 258
DB 251 KLRKYRSHDDQMIKCLSLISYSATIGGLTTIIGTSTSLIFLHFNNQYPA-AEVVNF 309
QY 259 ASWFGFAPPTMIILLLAWLQVLFGLVFNFRKNFGFGEGBE-ERKQAAFOVIKTOYRL 317
DB 310 GTWFLSPFISLIMLVSVFWMHVLFLGCNFKETCSLKKKTKREQLSEKRIQEEYEKL 369
QY 318 GPMSPAECTVTVLVLLVLMVTFREPFGPPGMDTVFANEKQSMASDGTVAIFISLV 377
DB 370 GDISYEMVTGFFILMTVLMVTFREPFGVPGWDS--FFEKKG--YRTDATVSVFLGFL 425
QY 378 IIPSKIP--GLMQDPKPKGLKAPPAILTWTKTVDNKPWNIVILLGGGFALAKGSSQSL 435
DB 426 LIPAKKPCPKGKNDGENQBSLGTESIITKDPQKTMPEIIVILVGGGYALASGSKS 485
QY 436 SEWLGDKLTPLQHPISPSATAVILCLLIAIIFTECTSNVATTTLFLPILASMAQICLHP 495
DB 486 STWIGNQMLSSLPWPVATVILACILSVISVTEFVNPNATITIFLILCSLSETMHPN 545
QY 496 VMLPCTLAASLAPLVPATPPNAIVFSFGGLKVSDMARAGFLNLIIGVLAITLSINSWS 555
DB 546 TLIPVTMCISFAVMLPVGNPPNAIVFSYGHQCIKDMVKAGLVNIGLVIMVAINTWG 605
QY 556 PIKLDTPPSWA 567
DB 606 SLFHLDTYPAWA 617
```

RESULT 7
US-09-556-916-10

; Sequence 10, Application US/09556916
; Patent No. 6548271

GENERAL INFORMATION:

; APPLICANT: Turner, Alex
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Sands, Arthur T.

; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins

; FILE REFERENCE: 8535-041-999

; CURRENT APPLICATION NUMBER: US/09/556,916

; CURRENT FILING DATE: 2000-04-21

; NUMBER OF SEQ ID NOS: 32

; SOFTWARE: FastSEQ for Windows Version 3.0

; SEQ ID NO 10

; LENGTH: 627

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-556-916-10

Query Match 41.0%; Score 1254; DB 4; Length 627;

Best Local Similarity 41.2%; Pred. No. 8.6e-119;

Matches 253; Conservative 124; Mismatches 173; Indels 64; Gaps 13;

QY 12 RYLVLCPLPFLPLPLIVQKEAYCAYSIILMALLMCTEALPLAVTALPVLFPPLMG 71

DB 11 RKLWVCPVLLPLPLVHPSESCAYVLTAVYVWSEAVPLGAAALVPAFLYPPFG 70

QY 72 IMDASEVCIEYFKDNTILFVGLMVAIAVEHNLHKLRIALQVLLIIGVRPALLLIGFMLV 131

DB 71 VLRSEVAAYEYFKNTLLVGVICVAAAVEKNLHKLRIALRMVLMAGAKPGMLLFCFCC 130

QY 132 TAFLSMISNTATTAMVPVIGHAVLEQLQSGK-KDVEGNNN-----PTF 175

DB 131 TTLSSMWSNTSTTAMVPVIEAVLQELVSADEQLVAGNSNTEAEPISLDVKNQPSL 190

QY 176 EL---QECPQKEVTKLD-----NQOP-----VSAPSEPR 202

DB 191 ELIFVNEEDRSNADLTTLMHENENLNGVPSITNPITKANQHKGKQHPSEQKQVLTFS-PR 249

QY 203 TQKT---QEHH--RFSQGLSLCICYSASIGGIATLTGTTNVLVQGVNSLFPQNGNV 256

DB 250 KQKLRKYSRHHQDMICKLSLSISYSATIGGLTTIIGTSLIFLEHFNQYPA-AEVV 308

QY 257 NFASWFGFAPTMIILLLAWLQVLGVNFRKNFGEGEE-ERKQAAFOVIKTOYR 315

DB 309 NFGTWLFSPISLIMLVSWFMWHLFGCNFKETCSLKKKTKRQLSEKRIQEBYE 368

QY 316 LLGPMSPAEKTVTVLLVLLVWFTREPFGPPGMDTVFANEKGQSMASDGTVAIFISLV 375

DB 369 KLGDISYPMVGTGPFILMTVLMFTREPFGVPGWDS--FPEKKG--YRTDATVSFVLGFL 424

QY 376 MFIIPSKIP--GLMQDPKPKGLKAPPAILTWKTVNDQMPNIVILLGGGALAGSQS 433

DB 425 LFLIPAKKCFCKNDGQNEHSLGTEIITWKDFQKTPWEIVLVGGVALAGSKSS 484

QY 434 GLSEWLGKLTPLQHPISATAVILCLLIAIIFTECTSNVATTLPLPLASMAQAICLHP 493

DB 485 GLSTWIGNQMLSSLPSPWAVTLLACILSVITFVSNPATITIFPLCLSLSETQHIN 544

QY 494 LYVMPCTLAASLAFMLPVATPPNAIVFSFGGLKVSDMARAGFLNIIIGVLAITLSINSW 553

DB 545 LYTLPVTWCISFAVMLPVGNPPNAIVFSYGHQCIKDMVKAGLVNVLGLVIMVAINTW 604

QY 554 SIPIKLDTPPSWA 567

DB 605 GVSLFHLDTYPAWA 618

RESULT 8

US-09-556-916-14

; Sequence 14, Application US/09556916

; Patent No. 6548271

GENERAL INFORMATION:

; APPLICANT: Turner, Alex
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Sands, Arthur T.

; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins

; FILE REFERENCE: 8535-041-999

; CURRENT APPLICATION NUMBER: US/09/556,916

; CURRENT FILING DATE: 2000-04-21

; NUMBER OF SEQ ID NOS: 32

; SOFTWARE: FastSEQ for Windows Version 3.0

; SEQ ID NO 14

; LENGTH: 626

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-556-916-14

Query Match 41.0%; Score 1253.5; DB 4; Length 626;

Best Local Similarity 40.8%; Pred. No. 9.7e-119;

Matches 250; Conservative 122; Mismatches 179; Indels 61; Gaps 11;

QY 12 RYLVLCPLPFLPLPLIVQKEAYCAYSIILMALLMCTEALPLAVTALPVLFPPLMG 71

DB 11 RKLWVCPVLLPLPLVHPSESCAYVLTAVYVWSEAVPLGAAALVPAFLYPPFG 70

QY 72 IMDASEVCIEYFKDNTILFVGLMVAIAVEHNLHKLRIALQVLLIIGVRPALLLIGFMLV 131

DB 71 VLRSEVAAYEYFKNTLLVGVICVAAAVEKNLHKLRIALRMVLMAGAKPGMLLFCFCC 130

QY 132 TAFLSMISNTATTAMVPVIGHAVLEQLQSGK-KDVEGNNN-----PTF 175

DB 131 TTLSSMWSNTSTTAMVPVIEAVLQELVSADEQLVAGNSNTEAEPISLDVKNQPSL 190

QY 176 EL---QECPQKEVTKLDNGQVPSAPS-----EPRTO 204

DB 191 ELIFVNEESNADLTTLMHENENLNGVPSITNPITKANQHKGKQHPSEQKQVLTFSPRKQ 250

QY 205 KT---QEHH--RFSQGLSLCICYSASIGGIATLTGTTNVLVQGVNSLFPQNGNVNF 258

DB 251 KLRKYSRHHQDMICKLSLSISYSATIGGLTTIIGTSLIFLEHFNQYPA-AEVNF 309

QY 259 ASWFGFAPTMIILLLAWLQVLGVNFRKNFGEGEE-ERKQAAFOVIKTOYRLL 317

DB 310 GTWELFSPISLIMLVSWFMWHLFGCNFKETCSLKKKTKRQLSEKRIQEBEYKL 369

QY 318 GPMSPAEKTVTVLLVLLVWFTREPFGPPGMDTVFANEKGQSMASDGTVAIFISLVNF 377

DB 370 GDISYPMVGTGPFILMTVLMFTREPFGVPGWDS--FPEKKG--YRTDATVSFVLGFL 425

QY 378 IIPSKIP--GLMQDPKPKGLKAPPAILTWKTVNDQMPNIVILLGGGALAGSQSGL 435

DB 426 LIPAKKCFCKNDGQNEHSLGTEIITWKDFQKTPWEIVLVGGVALAGSKSSGL 485

QY 436 SEWLGKLTPLQHPISATAVILCLLIAIIFTECTSNVATTLPLPLASMAQAICLHP 495

DB 486 STWIGNQMLSSLPSPWAVTLLACILSVITFVSNPATITIFPLCLSLSETLHINPLY 545

QY 496 VMLPCTLAASLAFMLPVATPPNAIVFSFGGLKVSDMARAGFLNIIIGVLAITLSINSWSI 555

DB 546 TLIPVTWCISFAVMLPVGNPPNAIVFSYGHQCIKDMVKAGLVNVLGLVIMVAINTW 605

QY 556 PIKLDTPPSWA 567

DB 606 SLFHLDTYPAWA 617

RESULT 9

US-09-556-916-4

; Sequence 4, Application US/09556916

; Patent No. 6548271

; GENERAL INFORMATION:

; APPLICANT: Turner, Alex

; APPLICANT: Zambrowicz, Brian
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins
; FILE REFERENCE: 8535-041-999
; CURRENT APPLICATION NUMBER: US/09/556,916
; CURRENT FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 627
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-556-916-4

Query Match 40.9%; Score 1252; DB 4; Length 627;
Best Local Similarity 41.2%; Pred. No. 1.4e-118;
Matches 253; Conservative 123; Mismatches 174; Indels 64; Gaps 13;
QY 12 RFVLIVLCPIFLPLPLIVQTKYCAYSIILMALLWCTEALPLAVTALFPVLPPLMG 71
Db 11 RKLVLVVCVPLLLPLPLVLPSPSEASCAVLIIVTAVYVWSEAVPLGAAALVPAFLYPPFG 70
QY 72 IMDASEVCIEYKDTNLFVGGMLVAIVAEHVNHLKRIALQVLLIIGVRPALLLGLFMLV 131
Db 71 VLRSNEVAAYEYKNTTLLLVGVICVAAAEKVNHLKRIALRMVLMAGAKPGMLLCFMCC 130
QY 132 TAPLSMWISNTATTAMVPVPIGHALEQLQGSK-KDVEGGNNN-----PTF 175
Db 131 TTLLSMWLSNTSTTAMVMPVEAVLQELVGADEQLVAGNSNTEEAEPISLDVKNSQPSL 190
QY 176 EL---QECPQKEVTKLD-----NQDP-----VSAPSEPR 202
Db 191 ELIFVNEEDRNADLTLMHNENLVPSITNPIKTANQHQKQHPQSEKQPQVLTFS-PR 249
QY 203 TQKT-----QEH--RFSQGLSLCICYSASIGGIATLTGTPNVLQGVNSLFPQNGNVV 256
Db 250 KQKLNRYRSHHDQMICKCLSLSLSYSATIGGLTIIGTSTLSIFLEHFNQYPA-AEVV 308
QY 257 NFASWFGFAPPTWIIILLLWMLQVLFLGVNFRKQFGFGEGBE-ERKQAAFOVIKTQYR 315
Db 309 NFGTWFLFSPFISLIMLVVSWFMHMLFLGCNFKETCSLSKKKTKREQLSEKRIQEEYE 368
QY 316 LGPMSFAEKTVTVLVLLVVLVFTREPFGPGMDTVFANEKQGSMDGTVAIFISLV 375
Db 369 KLGDISPEVMVTGFFILMTVLVFTREPFGVPGWDS--FPEKKG--YRTDATVSVFLGFL 424
QY 376 MFIPSKIP--GLMQDPKPKGLKAPPAILTWTVDNPKMPNIVILGGGFALAKGSEQS 433
Db 425 LFLIPAKKPCFGKKNQDGENQHSLSGTPEIITWKDFQKTPMPEIIVLGGGYALASGSKS 484
QY 434 GLSEWGDKLTPLQHPSPSATVILCLLIAIFTECTSNVATTLTFLPILASMAQAICLHP 493
Db 485 GLSTWIGNQLSLSLSPWAVTLTACILVSIIVTEFVSNPATITIFLPILCSLSETOHNP 544
QY 494 LYVLPCTLAASLAFMLPVATPNAIVFSPGGLKVSDMARAGFLNLIIGVLAITLSINSW 553
Db 545 LYLIPVTCISFAVMLPVGNPNNAIVFSYGHCOIKDMVKAGLVNVLGVVAVNTW 604
QY 554 SIPIFKLDTFPPSWA 567
Db 605 GVSLFHLDTYPAWA 618

RESULT 10
US-09-556-916-22
; Sequence 22, Application US/09556916
; Patent No. 6548271
; GENERAL INFORMATION:
; APPLICANT: Turner, Alex
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Nehls, Michael

; APPLICANT: Friedrich, Glenn
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins
; FILE REFERENCE: 8535-041-999
; CURRENT APPLICATION NUMBER: US/09/556,916
; CURRENT FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 22
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-556-916-22

Query Match 40.9%; Score 1250.5; DB 4; Length 626;
Best Local Similarity 40.8%; Pred. No. 2e-118;
Matches 250; Conservative 122; Mismatches 179; Indels 61; Gaps 11;
QY 12 RFVLIVLCPIFLPLPLIVQTKYCAYSIILMALLWCTEALPLAVTALFPVLPPLMG 71
Db 11 RKLVLVVCVPLLLPLPLVLPSPSEASCAVLIIVTAVYVWSEAVPLGAAALVPAFLYPPFG 70
QY 72 IMDASEVCIEYKDTNLFVGGMLVAIVAEHVNHLKRIALQVLLIIGVRPALLLGLFMLV 131
Db 71 VLRSNEVAAYEYKNTTLLLVGVICVAAAEKVNHLKRIALRMVLMAGAKPGMLLCFMCC 130
QY 132 TAPLSMWISNTATTAMVPVPIGHALEQLQGSK-KDVEGGNNN-----PTF 175
Db 131 TTLLSMWLSNTSTTAMVMPVEAVLQELVGADEQLVAGNSNTEEAEPISLDVKNSQPSL 190
QY 176 EL---QECPQKEVTKLDNGQPVSAFS-----EPRTQ 204
Db 191 ELIFVNEESNADLTLMHNENLVPSITNPIKTANQHQKQHPQSEKQPQVLTFSPRQ 250
QY 205 KT-----QEH--RFSQGLSLCICYSASIGGIATLTGTPNVLQGVNSLFPQNGNVNP 258
Db 251 KLNRYRSHHDQMICKCLSLSLSYSATIGGLTIIGTSTLSIFLEHFNQYPA-AEVVNP 309
QY 259 ASWFGFAPPTWIIILLLWMLQVLFLGVNFRKQFGFGEGBE-ERKQAAFOVIKTQYRLL 317
Db 310 GTWFLFSPFISLIMLVVSWFMHMLFLGCNFKETCSLSKKKTKREQLSEKRIQEEYEKL 369
QY 318 GPMSFAEKTVTVLVLLVVLVFTREPFGPGMDTVFANEKQGSMDGTVAIFISLVNP 377
Db 370 GDISPEVMVTGFFILMTVLVFTREPFGVPGWDS--FPEKKG--YRTDATVSVFLGFLF 425
QY 378 ITPSKIP--GLMQDPKPKGLKAPPAILTWTVDNPKMPNIVILGGGFALAKGSEQSGL 435
Db 426 LIPAKKPCFGKKNQDGENQHSLSGTESIITWKDFQKTPMPEIIVLGGGYALASGSKSGL 485
QY 436 SEWLGDKLTPLQHPSPSATVILCLLIAIFTECTSNVATTLTFLPILASMAQAICLHP 495
Db 486 STWIGNQLSLSLSPWAVTLTACILVSIIVTEFVSNPATITIFLPILCSLSETOHNP 545
QY 496 VMLPCTLAASLAFMLPVATPNAIVFSPGGLKVSDMARAGFLNLIIGVLAITLSINSW 555
Db 546 TLIPVTCISFAVMLPVGNPNNAIVFSYGHCOIKDMVKAGLVNVLGVVAVNTW 605
QY 556 PIPIFKLDTFPPSWA 567
Db 606 SLFHLDTYPAWA 617

RESULT 11
US-09-556-916-16
; Sequence 16, Application US/09556916
; Patent No. 6548271
; GENERAL INFORMATION:
; APPLICANT: Turner, Alex
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Sands, Arthur T.

; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins
; FILE REFERENCE: 8535-041-999
; CURRENT APPLICATION NUMBER: US/09/556,916
; CURRENT FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-556-916-16

Query Match 40.8%; Score 1248.5; DB 4; Length 626;
Best Local Similarity 40.8%; Pred. No. 3.le-118;
Matches 250; Conservative 121; Mismatches 180; Indels 61; Gaps 11;
QY 12 RFLVILVCLPIPLPLPLVQTEAKYCAYSITLMALLMCTEALPLAVTALFPIVPLPLMG 71
DB 11 RKLLLVVCPVLLPLVLPVHPSEASCAYLVITAVYVSEAVPLGAAALVPAFLYPPFG 70
QY 72 IMDASEVCIEYFKDNTNLFVGGMLVAIAVEHNLHKLIALQVLLIIGVRPALLLGLFMLV 131
DB 71 VLRSNEVAAYEYFKNITLLLVGVICVAAAVEKNLHKLIALRMVLMAGAKPGMLLLCFMCC 130
QY 132 TAFLSMWISNTATTAMVPVIGHAVLEOQSGK-KDVEGGNN-----PTF 175
DB 131 TLLSMWLSNTSTTAMVPIVEAVLQELVSAEDEQLVAGNSNTEEAEPISLDVKNQSPSL 190
QY 176 EL---QEBCPQKEVTKLDNGQPVSA-----VSAPSEPR 202
DB 191 ELIFNEESNADLTTLNHNENLVPSITNPIKTAHQHGKQHPQVLTSPRKQ 250
QY 205 KT-----QEHH--RFSQGLSLCICYSASIGGIATLTGTPNLVLOGVNSLFPQNGVNF 258
DB 251 KLNRYRSHDDQWICKLSLSISYSATIGLTTIGTSTLSIFLEHFNQYPA-AEVNF 309
QY 259 ASWFGFAPFTMIILLALLWLQVFLGVNFRKNFGFEGEE-ERKQAAFOVKTQYRL 317
DB 310 GTWFLFSPISLIMLVSMFWMHFLPGCNFKETCSLSKKKTKREQLSEKRIQBEYEKL 369
QY 318 GPMSPAEKTVTVLVLLVLTREPGFPFGDVTVPANEKQSGMASDGTVAIFISLWPF 377
DB 370 GDISPEWMTGFFILMTVLTREPGFVPGWDS--FPEKKG--YRTDATSVFLGFLF 425
QY 378 IIPSKIP--GLMQDPKPKGLKAPAILTWKTVNDKMPNIVILLGGGFALAKSGESQGL 435
DB 426 LIPAKKPCFGKNDGENQHSLSGTPEIITWKDFQKTPWEIIVLGGGVYALASGSKSGL 485
QY 436 SEWLGDKLTPLQHIPPSATAVILCLLIAIETECTSNVATTTPLPLASMAQICLHP 495
DB 486 STWIGNQMLSSLSLPPWAVTLLACILSVITFEVSNPATITITPLILCSLSETHINPLY 545
QY 496 VMLPCTLAASLAPMLPVATPPNAIVPSFGLKVSMDARAGFLNIGVLAITLSINSWI 555
DB 546 TLIPVTMCISFAMLPVGNPPNAIVPSYCHQIKDMVKAGLVNIGLVIMVAINTWGV 605
QY 556 PIFKLDTPPSWA 567
DB 606 SLFHLDTYPAWA 617

RESULT 12
US-09-556-916-12
; Sequence 12, Application US/09556916
; Patent No. 6548271
; GENERAL INFORMATION:
; APPLICANT: Turner, Alex
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins
; FILE REFERENCE: 8535-041-999

; CURRENT APPLICATION NUMBER: US/09/556,916
; CURRENT FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 581
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-556-916-12

Query Match 35.9%; Score 1097; DB 4; Length 581;
Best Local Similarity 37.6%; Pred. No. 8e-103;
Matches 231; Conservative 114; Mismatches 159; Indels 110; Gaps 14;
QY 12 RFLVILVCLPIPLPLPLVQTEAKYCAYSITLMALLMCTEALPLAVTALFPIVPLPLMG 71
DB 11 RKLLLVVCPVLLPLVLPVHPSEASCAYLVITAVYVSEAVPLGAAALVPAFLYPPFG 70
QY 72 IMDASEVCIEYFKDNTNLFVGGMLVAIAVEHNLHKLIALQVLLIIGVRPALLLGLFMLV 131
DB 71 VLRSNEVAAYEYFKNITLLLVGVICVAAAVEKNLHKLIALRMVLMAGAKPGMLLLCFMCC 130
QY 132 TAFLSMWISNTATTAMVPVIGHAVLEOQSGK-KDVEGGNN-----PTF 175
DB 131 TLLSMWLSNTSTTAMVPIVEAVLQELVSAEDEQLVAGNSNTEEAEPISLDVKNQSPSL 190
QY 176 EL---QEBCPQKEVTKLDNGQPVSA-----VSAPSEPR 202
DB 191 ELIFNEESNADLTTLNHNENLVPSITNPIKTAHQHGKQHPQVLTSPRKQ 250
QY 205 KT-----QEHH--RFSQGLSLCICYSASIGGIATLTGTPNLVLOGVNSLFPQNGVNF 258
DB 251 KLNRYRSHDDQWICKLSLSISYSATIGLTTIGTSTLSIFLEHFNQYPA-AEVNF 309
QY 259 ASWFGFAPFTMIILLALLWLQVFLGVNFRKNFGFEGEE-ERKQAAFOVKTQYRL 317
DB 310 GTWFLFSPISLIMLVSMFWMHFLPGCNFKETCSLSKKKTKREQLSEKRIQBEYEKL 369
QY 318 GPMSPAEKTVTVLVLLVLTREPGFPFGDVTVPANEKQSGMASDGTVAIFISLWPF 377
DB 370 GDISPEWMTGFFILMTVLTREPGFVPGWDS--FPEKKG--YRTDATSVFLGFLF 425
QY 378 IIPSKIP--GLMQDPKPKGLKAPAILTWKTVNDKMPNIVILLGGGFALAKSGESQGL 435
DB 426 LIPAKKPCFGKNDGENQHSLSGTPEIITWKDFQKTPWEIIVLGGGVYALASGSKSGL 485
QY 436 SEWLGDKLTPLQHIPPSATAVILCLLIAIETECTSNVATTTPLPLASMAQICLHP 493
DB 485 GLTWIGNQMLSSLSLPPWAVTLLACILSVITFEVSNPATITITPLILCSLSETHINPLY 535
QY 494 LYVMLPCTLAASLAPMLPVATPPNAIVPSFGLKVSMDARAGFLNIGVLAITLSINSW 553
DB 536 -----LVKAGLVNIGLVIMVAINTW 558
QY 554 SIPIFKLDTPPSWA 567
DB 559 GVSFLHDTYPAWA 572

RESULT 13
US-09-556-916-6
; Sequence 6, Application US/09556916
; Patent No. 6548271
; GENERAL INFORMATION:
; APPLICANT: Turner, Alex
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins
; FILE REFERENCE: 8535-041-999
; CURRENT APPLICATION NUMBER: US/09/556,916
; CURRENT FILING DATE: 2000-04-21

; NUMBER OF SEQ ID NOS: 32
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 6
 ; LENGTH: 581
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-556-916-6

Query Match 35.8%; Score 1095; DB 4; Length 581;
 Best Local Similarity 37.6%; Pred. No. 1.3e-102;
 Matches 231; Conservative 113; Mismatches 160; Indels 110; Gaps 14;
 QY 12 RYLVLCPIELPLPLIVQTEKAYCAYSIIMALLNCTEALPLAVTALPFIPLPLMG 71
 Db 11 RKLIVVCPVLLPLPLVLPSPSEASAYVLIIVAVYVSEAVPLGNAALVPAFLPFFG 70
 QY 72 IMDASEVCIFYKDTNLFVGGMLVAIAVEHNLHRIALQVLLIIGVRPALLLGFMLV 131
 Db 71 VLSRNEVAEYFNKNTLLVGVICVAAVEKNLHRIALRMVLMAGAKPGMLLCFMCC 130
 QY 132 TAFSLMWISNTATTAMVPIGHAVLEQLOQSK-KOVEGGNN-PTF 175
 Db 131 TTLSMWLSNTSTTAMVMPIVEAVLQELVSAEDQLVAGNSNTEEAEPISLDVKNQPSL 190
 QY 176 EL---QECPQKEVTKLD---NGQP-----VSAPSEPR 202
 Db 191 ELIFVNEEDRNADLTLLMHNENLVPSITNPIKTANQHQKQHPQSEKPVLTSP-PR 249
 QY 203 TQKT----QEHH--RFSQGLSLCICYSASIGGIATLTGTTNPLVLOQVNSLPFGQNVV 256
 Db 250 KQLNRKYRSHDDQMIKCLSLSSISYSATIGLTIIGTSLIFLHFNNQYPA-AEVV 308
 QY 257 NFASWGFAPPTMIILLLLAWLQVLFVGNFRKNFGFGEGB-ERKQAAFOVKTQYR 315
 Db 309 NFGWFLSPFISLIMLVSNFMHWFPGCNFKETCSLSKKKTKRQSEKRIQEEYE 368
 QY 316 LLGPMSPAEKTVTVLVLLVLTFTREPFGPGMGDTVFANEGKOSMASDGTVAIFSLV 375
 Db 369 KLGDISPEWVTGFFILMTVLTFTREPFGVPGWDS--FPEKKG--YRTDATVSVFLGL 424
 QY 376 MFIIPSKIP--GLMDPKPKGKAPPAITLTKTNDKMPNIVILGGGFALAKSGEQS 433
 Db 425 LFLIPAKKPCFGKKNKDNENQHSLSGTETIITKDFQKTPMPEIIVLGGGYALASGSKS 484
 QY 434 GLSEWGLKLTPLQHIPSPATAVILCLLIAIFTECTSNVATTTLFLPLASMAQAICLHP 493
 Db 485 GLSTWIGNQMLSLSLPPWAVTLACILVSVITFVSNPATITIFLPLICS-----535
 QY 494 LYVMPCTLAASLAFMPLPVATPPNAIVFSFGLKVSMDARAGFLNIIGVLAITLSINSW 553
 Db 536 -----LVKAGLVNVIGLVIVMVAINTW 558

RESULT 14
 US-09-556-916-24
 ; Sequence 24, Application US/09556916
 ; Patent No. 6548271
 ; GENERAL INFORMATION:
 ; APPLICANT: Turner, Alex
 ; APPLICANT: Zambrowicz, Brian
 ; APPLICANT: Nehls, Michael
 ; APPLICANT: Friedrich, Glenn
 ; APPLICANT: Sands, Arthur T.
 ; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins
 ; FILE REFERENCE: 8535-041-999
 ; CURRENT APPLICATION NUMBER: US/09/556,916
 ; CURRENT FILING DATE: 2000-04-21
 ; NUMBER OF SEQ ID NOS: 32
 ; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 24
 ; LENGTH: 580
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-556-916-24

Query Match 35.8%; Score 1093.5; DB 4; Length 580;
 Best Local Similarity 37.3%; Pred. No. 1.8e-102;
 Matches 228; Conservative 112; Mismatches 165; Indels 107; Gaps 12;
 QY 12 RYLVLCPIELPLPLIVQTEKAYCAYSIIMALLNCTEALPLAVTALPFIPLPLMG 71
 Db 11 RKLIVVCPVLLPLPLVLPSPSEASAYVLIIVAVYVSEAVPLGNAALVPAFLPFFG 70
 QY 72 IMDASEVCIFYKDTNLFVGGMLVAIAVEHNLHRIALQVLLIIGVRPALLLGFMLV 131
 Db 71 VLSRNEVAEYFNKNTLLVGVICVAAVEKNLHRIALRMVLMAGAKPGMLLCFMCC 130
 QY 132 TAFSLMWISNTATTAMVPIGHAVLEQLOQSK-KOVEGGNN-PTF 175
 Db 131 TTLSMWLSNTSTTAMVMPIVEAVLQELVSAEDQLVAGNSNTEEAEPISLDVKNQPSL 190
 QY 176 EL---QECPQKEVTKLDNGQPVSAPS-----EPRTQ 204
 Db 191 ELIFVNEESNADLTLLMHNENLVPSITNPIKTANQHQKQHPQSEKPVLTSPRKQ 250
 QY 205 KT----QEHH--RFSQGLSLCICYSASIGGIATLTGTTNPLVLOQVNSLPFGQNVV 258
 Db 251 KQLNRKYRSHDDQMIKCLSLSSISYSATIGLTIIGTSLIFLHFNNQYPA-AEVV 309
 QY 259 ASWFGFAPPTMIILLLLAWLQVLFVGNFRKNFGFGEGB-ERKQAAFOVKTQYRLL 317
 Db 310 GTWFLSPFISLIMLVSNFMHWFPGCNFKETCSLSKKKTKRQSEKRIQEEYEKL 369
 QY 318 GPMSPAEKTVTVLVLLVLTFTREPFGPGMGDTVFANEGKOSMASDGTVAIFSLV 377
 Db 370 GDISYPEWVTGFFILMTVLTFTREPFGVPGWDS--FPEKKG--YRTDATVSVFLGLLF 425
 QY 378 IIPSKIP--GLMDPKPKGKAPPAITLTKTNDKMPNIVILGGGFALAKSGEQSGL 435
 Db 426 LIPAKKPCFGKKNKDNENQHSLSGTETIITKDFQKTPMPEIIVLGGGYALASGSKS 485
 QY 436 SEWGLKLTPLQHIPSPATAVILCLLIAIFTECTSNVATTTLFLPLASMAQAICLHP 495
 Db 486 STWIGNQMLSLSLPPWAVTLACILVSVITFVSNPATITIFLPLICS-----534
 QY 496 VMLPCTLAASLAFMPLPVATPPNAIVFSFGLKVSMDARAGFLNIIGVLAITLSINSW 555
 Db 535 -----LVKAGLVNVIGLVIVMVAINTW 559

RESULT 15
 US-09-556-916-18
 ; Sequence 18, Application US/09556916
 ; Patent No. 6548271
 ; GENERAL INFORMATION:
 ; APPLICANT: Turner, Alex
 ; APPLICANT: Zambrowicz, Brian
 ; APPLICANT: Nehls, Michael
 ; APPLICANT: Friedrich, Glenn
 ; APPLICANT: Sands, Arthur T.
 ; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins
 ; FILE REFERENCE: 8535-041-999
 ; CURRENT APPLICATION NUMBER: US/09/556,916
 ; CURRENT FILING DATE: 2000-04-21
 ; NUMBER OF SEQ ID NOS: 32
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 18
 ; LENGTH: 580

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-556-916-18

Query Match      35.7%; Score 1091.5; DB 4; Length 580;
Best Local Similarity 37.3%; Pred. No. 2.9e-102;
Matches 228; Conservative 111; Mismatches 166; Indels 107; Gaps 12;

QY 12 RFYLVLCPLPFIPLPLIVQTKAYCAYSIIIMALLWCTEALPLAVTALFPVLPPLMG 71
Db 11 RKLLLVVCPVLLPLPLVLPSPSEASCAYLVITAVYVSEAVPLGAAALVDAFLYPPFG 70

QY 72 INDASEVCIEYKPDNTILFVGGIMVAIVHEHNLHKRIALQVLLIIGVVPALLILGFMV 131
Db 71 VLRSNEVAARYFKNTITLLVGVTCVAAAVKRWNLHKRIALRMVLMAGAKPGMLLLCFMCC 130

QY 132 TAPLSMWISNTATTAMVPIGHAVLEQLQSGK-KVEGGGNNN-----PTF 175
Db 131 TTLLSMWLSNTSTTAMVPIVEAVLQELVSADEQLVAGNSNTEEAEPISLDVKNSQPSL 190

QY 176 EL---QEECPQKEVTKLDNGQPVSAFS-----EPRTQ 204
Db 191 ELIFVNEESNADLTTLHMENLNGVPSITNPIKTANHQKQKQHPQKQKPVLTSPRKQ 250

QY 205 KT-----QEHK--RFSQGLSLCICYSASIGGIATLTGTTNVLVQQGVNSLFPQNGNVNF 258
Db 251 KLNRYRSHHQWICKCLSLISYSATIGGLTIIGTSTSLIFLHFNNQYPA-AEVVNF 309

QY 259 ASWFGFAPRTMILLALLMLQVLPLGVNFRKNFGFGGBE-ERKQAAFOVIKTQYRL 317
Db 310 GTWFLFSFISLIMLVVSWFMHVLFLGCNFKETCSLSKKKTKKREQUSEKRIQBEYEKL 369

QY 318 GPMFAEKTVTTVLVLLVVLWFTREPPFPFGMDTVFANEKGQSMASDGTVAIFISLVNF 377
Db 370 GDISYPEMYTGFFILMTVLMFTREPGFVPGWDS--FPEKKG--YETDATSVFLGFLLF 425

QY 378 IIPSKIP--GLMQDPKPKGLKAPAILTWKTVNDKMPNIVILGGGFALAKGSEQSGL 435
Db 426 LIPAKKPCFCGKNDGENQSHSLGTEPIITWKDFQKTPWEIVILVGGGYALASGSKSGL 485

QY 436 SEMLGDKLTPLQHIIPSATAVILCLLJALFTECTSNVATTTLPPLILASMAQAICLHPLY 495
Db 486 STWIGNQMLSSLSPWPWATLACILSVITFEVSNPATITIFLPILCS-----534

QY 496 VMLPCTLAASLAFMLPVATPPNAIVFSGGLKVSDMARAGFLINIIGVLAITLSINWSI 555
Db 535 -----LVKAGLGVNVLGVVNVVAINTWGV 559

QY 556 PIFKLDTPPSWA 567
Db 560 SLFHLDTYPWA 571
```

Search completed: June 30, 2005, 09:02:16
Job time : 25.5692 secs

THIS PAGE IS BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2005, 08:59:01 ; Search time 76.171 Seconds
(without alignments)
2973.073 Million cell updates

Title: US-10-017-479A-4

Perfect s

Sequence: 1 MATCWPALWAYRFYLVCL.....HSNTSQCLLNPSNSTVPGGL 587

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1722976 seqs. 385795295

Total number of hits satisfying chosen parameters: 1722976

Minimum DB seq length: 0

Maximum DB seq length:	0
Maximum DB seq length:	2000000000

Post-processing: Minimum Match 0%

Maximum Match 100
Listing first 100

Database : Published Applications

```
1: /cgn2_6/ptodab
2: /cgn2_5/ptodab
```

3: /cgn2_6/ptodat

5: /cgn2_6/ptodat

7: /cgn2_6/ptodat

9: /cgn2_6/ptodat

11: /cgn2_6/ptoda

13: /cgn2_6/ptoda

11: /cgn2_6/ptoda
15: /cgn2_6/ptoda

```
16: /cgn2_8/ptoda
17: /cgn2_6/ptoda
```

18: /cgnz_6/ptoda
19: /cgnz_6/ptoda

```
20: /cgn2_6/ptoda
21: /cgn2_6/ptoda
```

22: /cgn2_6/ptoda

Pred. No. is the number of r

and is derived by analysis of

Result	Query
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91
92	92
93	93
94	94
95	95
96	96
97	97
98	98
99	99
100	100

[illegible]

2	3058	100.0	587	14
---	------	-------	-----	----

4	2392	78.2	592	14
4	2392	78.2	592	14

	6	6	9	9	U
1971.5	64.5	619	9	U	

7	1371.3	64.3	613	14
8	1829	59.8	581	17

3	1028	33.2	616	13
10	1628	53.2	616	15

TT 879T 53.2 919 15

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	3058	100.0	587	13	US-10-173-519-5	Sequence 5, Appli
2	3058	100.0	587	14	US-10-167-994-12	Sequence 12, Appli
3	3058	100.0	587	17	US-10-718-359-13	Sequence 13, Appli
4	2392	78.2	592	14	US-10-167-994-3	Sequence 3, Appli
5	2392	78.2	592	14	US-10-167-994-11	Sequence 11, Appli
6	1971.5	64.5	619	9	US-09-729-094-4	Sequence 4, Appli
7	1971.5	64.5	619	14	US-10-435-631-4	Sequence 4, Appli
8	1829	59.8	581	17	US-10-718-359-12	Sequence 12, Appli
9	1628	53.2	616	15	US-10-092-900A-46	Sequence 46, Appli
10	1628	53.2	616	15	US-10-092-900A-48	Sequence 48, Appli
11	1628	53.2	616	15	US-10-403-161-68	Sequence 68, Appli

85 358.5 11.7 162 11 US-09-833-245-1948 Sequence 1948, Ap
86 350.5 11.5 162 11 US-09-833-245-1945 Sequence 1945, Ap
87 336 11.0 1006 17 US-10-741-849-7011 Sequence 7011, Ap
88 335.5 11.0 499 14 US-10-238-075-500 Sequence 500, App
89 327 10.7 259 16 US-10-739-930-9458 Sequence 9458, Ap
90 311.5 10.2 300 16 US-10-767-701-46317 Sequence 46317, A
91 299 9.8 457 15 US-10-369-493-4366 Sequence 4266, Ap
92 289 9.5 78 9 US-09-864-761-34352 Sequence 34252, A
93 289 9.5 78 14 US-10-029-386-28038 Sequence 28038, A
94 289 9.5 174 16 US-10-425-115-335649 Sequence 335649, A
95 285 9.3 163 15 US-10-424-599-194881 Sequence 184881, A
96 258.5 8.5 135 15 US-10-276-774-1592 Sequence 1592, Ap
97 256.5 8.4 487 15 US-10-369-493-749 Sequence 749, App
98 240.5 7.9 164 15 US-10-424-599-205152 Sequence 205152, A
99 235.5 7.7 487 15 US-10-369-493-23559 Sequence 23559, A
100 208 6.8 94 9 US-09-864-761-43642 Sequence 43642, A

ALIGNMENTS

RESULT 1
US-10-173-519-5
; Sequence 5, Application US/10173519
; Publication No. US2002019582A1
; GENERAL INFORMATION:
; APPLICANT: Curtiss, Rory A.J.
; APPLICANT: Millennium Pharmaceuticals, Inc.
; TITLE OF INVENTION: 69624, A Human Transporter Family Member
; TITLE OF INVENTION: and Uses Thereof
; FILE REFERENCE: MP101-098P1RNM
; CURRENT APPLICATION NUMBER: US/10/173,519
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 60/298,970
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 587
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-173-519-5

Query Match 100.0%; Score 3058; DB 13; Length 587;
Best Local Similarity 100.0%; Pred. No. 1.4e-280;
Matches 587; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATCPALWAYRYFLIVLCLPIFLPLPIVOTKEAYCAYSIILMALLMCTEALPLAVTA 60
DB 1 MATCPALWAYRYFLIVLCLPIFLPLPIVOTKEAYCAYSIILMALLMCTEALPLAVTA 60
QY 61 LFPDIVLPLMGINDASEVCIEYFKDTNLLFVGLMVAIAVEHWNHAKRIALQVLLIIGVR 120
DB 61 LFPDIVLPLMGINDASEVCIEYFKDTNLLFVGLMVAIAVEHWNHAKRIALQVLLIIGVR 120
QY 121 PALLLGFMVLVTAFLSMWISNTATTAMVPIGHAVLEQLQSKKDVGGNNNPTFELQEE 180
DB 121 PALLLGFMVLVTAFLSMWISNTATTAMVPIGHAVLEQLQSKKDVGGNNNPTFELQEE 180
QY 181 CPQKEVTKLDNGQPVSAPESEPTQKTOEHRRFSQGLSLCICYSASIGGIATLTGTTNVL 240
DB 181 CPQKEVTKLDNGQPVSAPESEPTQKTOEHRRFSQGLSLCICYSASIGGIATLTGTTNVL 240
QY 241 LOGQVNSLFPQNGVNVNFASFAGFPTMIILLLAWLQVLFVGNFRKNFGFGGEE 300
DB 241 LOGQVNSLFPQNGVNVNFASFAGFPTMIILLLAWLQVLFVGNFRKNFGFGGEE 300
QY 301 ERKQAAFOVTKTOYRLGPMGSAEKTIVLVLVLLVMTREPFGFGWGDVTFANEKGQ 360
DB 301 ERKQAAFOVTKTOYRLGPMGSAEKTIVLVLVLLVMTREPFGFGWGDVTFANEKGQ 360
QY 361 SMASDGTVAIFISLVNFIIPSKIPGLMQDPKKPGKLGKAPPAILTWKTVNDKMPNIVILL 420

DB 361 SMASDGTVAIFISLVNFIIPSKIPGLMQDPKKPGKLGKAPPAILTWKTVNDKMPNIVILL 420
QY 421 GGGFALAKGSEQSLSEWLGDKLTPLQHLPPSATAVILCLLIAIETECTSNVATTLFLP 480
DB 421 GGGFALAKGSEQSLSEWLGDKLTPLQHLPPSATAVILCLLIAIETECTSNVATTLFLP 480
QY 481 ILASMAQAICLHPLYVMLPCTLAASLAFMLPVATPPNAIVFSFGGLKVSDMARAGFLNI 540
DB 481 ILASMAQAICLHPLYVMLPCTLAASLAFMLPVATPPNAIVFSFGGLKVSDMARAGFLNI 540
QY 541 IGVLAITISINSWSPIFKLDTFPPSAHSNTSQCLLNPSNSTVPGGL 587
DB 541 IGVLAITISINSWSPIFKLDTFPPSAHSNTSQCLLNPSNSTVPGGL 587
RESULT 2
US-10-167-994-12
; Sequence 12, Application US/10167994
; Publication No. US20030082647A1
; GENERAL INFORMATION:
; APPLICANT: Reenan, Robert A.
; APPLICANT: Rogina, Blanka
; APPLICANT: Helfand, Stephen L.
; TITLE OF INVENTION: TRANSPORTER PROTEIN
; FILE REFERENCE: 13407-013001
; CURRENT APPLICATION NUMBER: US/10/167,994
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/255,013
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 10/017,479
; PRIOR FILING DATE: 2001-12-12
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 587
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-167-994-12

Query Match 100.0%; Score 3058; DB 14; Length 587;
Best Local Similarity 100.0%; Pred. No. 1.4e-280;
Matches 587; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATCPALWAYRYFLIVLCLPIFLPLPIVOTKEAYCAYSIILMALLMCTEALPLAVTA 60
DB 1 MATCPALWAYRYFLIVLCLPIFLPLPIVOTKEAYCAYSIILMALLMCTEALPLAVTA 60
QY 61 LFPDIVLPLMGINDASEVCIEYFKDTNLLFVGLMVAIAVEHWNHAKRIALQVLLIIGVR 120
DB 61 LFPDIVLPLMGINDASEVCIEYFKDTNLLFVGLMVAIAVEHWNHAKRIALQVLLIIGVR 120
QY 121 PALLLGFMVLVTAFLSMWISNTATTAMVPIGHAVLEQLQSKKDVGGNNNPTFELQEE 180
DB 121 PALLLGFMVLVTAFLSMWISNTATTAMVPIGHAVLEQLQSKKDVGGNNNPTFELQEE 180
QY 181 CPQKEVTKLDNGQPVSAPESEPTQKTOEHRRFSQGLSLCICYSASIGGIATLTGTTNVL 240
DB 181 CPQKEVTKLDNGQPVSAPESEPTQKTOEHRRFSQGLSLCICYSASIGGIATLTGTTNVL 240
QY 241 LOGQVNSLFPQNGVNVNFASFAGFPTMIILLLAWLQVLFVGNFRKNFGFGGEE 300
DB 241 LOGQVNSLFPQNGVNVNFASFAGFPTMIILLLAWLQVLFVGNFRKNFGFGGEE 300
QY 301 ERKQAAFOVTKTOYRLGPMGSAEKTIVLVLVLLVMTREPFGFGWGDVTFANEKGQ 360
DB 301 ERKQAAFOVTKTOYRLGPMGSAEKTIVLVLVLLVMTREPFGFGWGDVTFANEKGQ 360
QY 361 SMASDGTVAIFISLVNFIIPSKIPGLMQDPKKPGKLGKAPPAILTWKTVNDKMPNIVILL 420
DB 361 SMASDGTVAIFISLVNFIIPSKIPGLMQDPKKPGKLGKAPPAILTWKTVNDKMPNIVILL 420
QY 421 GGGFALAKGSEQSLSEWLGDKLTPLQHLPPSATAVILCLLIAIETECTSNVATTLFLP 480

Db 421 GGGFALAKGSEGLSEWLGDKLTPLQHIPPPSATAVILCLLIAIFTECTSNVATTTLFLP 480
Qy 481 ILASMAQAICLHPLYVMLPCTLAASLAFMLPVATPPNAIVFSGGLKVSMDMARAGFLNI 540
Db 481 ILASMAQAICLHPLYVMLPCTLAASLAFMLPVATPPNAIVFSGGLKVSMDMARAGFLNI 540
Qy 541 IGVLAITLSINSWSIPFKLDTFPPSWAHNSNTSQCLLNPSNSTVPGGL 587
Db 541 IGVLAITLSINSWSIPFKLDTFPPSWAHNSNTSQCLLNPSNSTVPGGL 587

RESULT 3
US-10-718-359-13
; Sequence 13, Application US/10718359
; Publication No. US2005095240A1
; GENERAL INFORMATION:
; APPLICANT: MEDICAL COLLEGE OF GEORGIA RESEARCH INSTITUTE, INC.
; TITLE OF INVENTION: NACT AS A TARGET FOR LIFESPAN EXPANSION AND WEIGHT REDUCTION
; FILE REFERENCE: 275.0008 0101
; CURRENT APPLICATION NUMBER: US/10/718,359
; PRIOR FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/428,469
; PRIOR FILING DATE: 2002-11-22
; PRIOR APPLICATION NUMBER: 60/459,441
; PRIOR FILING DATE: 2003-04-01
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 13
; LENGTH: 587
; TYPE: PRT
; ORGANISM: rat NaDC1
US-10-718-359-13

Query Match 100.0%; Score 3058; DB 17; Length 587;
Best Local Similarity 100.0%; Pred No. 1.4e-280;
Matches 587; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATCPALWAYRFLYLVLCLPIFLPLPLIVOTKEAYCAYSIIIMALLMCTEALPLAVTA 60
Db 1 MATCPALWAYRFLYLVLCLPIFLPLPLIVOTKEAYCAYSIIIMALLMCTEALPLAVTA 60
Qy 61 LPPIVLPFLMGIMDASEVCIIEYFKDTNILFVGLGMVAIAVEHNLHKRIALQVLLIIGVR 120
Db 61 LPPIVLPFLMGIMDASEVCIIEYFKDTNILFVGLGMVAIAVEHNLHKRIALQVLLIIGVR 120
Qy 121 PALLLGFMLVTAFLSMWISNTATTAMVPIGHAVLEQLOQSKKDVGGNNPTFELOEE 180
Db 121 PALLLGFMLVTAFLSMWISNTATTAMVPIGHAVLEQLOQSKKDVGGNNPTFELOEE 180
Qy 181 CPQKEVTKLDNGQVPSAPSEPRTKTQEHHRFSQGLSLCICYSASIGGIATLTGTPNLV 240
Db 181 CPQKEVTKLDNGQVPSAPSEPRTKTQEHHRFSQGLSLCICYSASIGGIATLTGTPNLV 240
Qy 241 LOGQVNSLPQNGNVNPFASWFGFAPPTMIILLALLAWLQVLFLGVNFRKNFGF 300
Db 241 LOGQVNSLPQNGNVNPFASWFGFAPPTMIILLALLAWLQVLFLGVNFRKNFGF 300
Qy 301 ERKQAAFOVIKQYRLGPMSPFAEKTVTVLVLLVVLVLTREPFGFGMDTVFANEKQ 360
Db 301 ERKQAAFOVIKQYRLGPMSPFAEKTVTVLVLLVVLVLTREPFGFGMDTVFANEKQ 360
Qy 361 SNASDGTVAIFSLVMFIIPSKIPGLMQDPKPKGLKAPPALLTWKTVNDKMPNIVILL 420
Db 361 SNASDGTVAIFSLVMFIIPSKIPGLMQDPKPKGLKAPPALLTWKTVNDKMPNIVILL 420
Qy 421 GGGFALAKGSEGLSEWLGDKLTPLQHIPPPSATAVILCLLIAIFTECTSNVATTTLFLP 480
Db 421 GGGFALAKGSEGLSEWLGDKLTPLQHIPPPSATAVILCLLIAIFTECTSNVATTTLFLP 480
Qy 481 ILASMAQAICLHPLYVMLPCTLAASLAFMLPVATPPNAIVFSGGLKVSMDMARAGFLNI 540
Db 481 ILASMAQAICLHPLYVMLPCTLAASLAFMLPVATPPNAIVFSGGLKVSMDMARAGFLNI 540

Qy 541 IGVLAITLSINSWSIPFKLDTFPPSWAHNSNTSQCLLNPSNSTVPGGL 587
Db 541 IGVLAITLSINSWSIPFKLDTFPPSWAHNSNTSQCLLNPSNSTVPGGL 587

RESULT 4
US-10-167-994-3
; Sequence 3, Application US/10167994
; Publication No. US20030082647A1
; GENERAL INFORMATION:
; APPLICANT: Reenan, Robert A.
; APPLICANT: Rogina, Blanka
; APPLICANT: Helfand, Stephen L.
; TITLE OF INVENTION: TRANSPORTER PROTEIN
; FILE REFERENCE: 13407-013001
; CURRENT APPLICATION NUMBER: US/10/167,994
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/255,013
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 10/017,479
; PRIOR FILING DATE: 2001-12-12
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 592
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-167-994-3

Query Match 78.2%; Score 2392; DB 14; Length 592;
Best Local Similarity 76.9%; Pred. No. 2e-217;
Matches 454; Conservative 63; Mismatches 67; Indels 6; Gaps 4;

Qy 1 MATCPALWAYRFLYLVLCLPIFLPLPLIVOTKEAYCAYSIIIMALLMCTEALPLAVTA 60
Db 1 MATCPALWAYRFLYLVLCLPIFLPLPLIVOTKEAYCAYSIIIMALLMCTEALPLAVTA 60
Qy 61 LPPIVLPFLMGIMDASEVCIIEYFKDTNILFVGLGMVAIAVEHNLHKRIALQVLLIIGVR 120
Db 61 LPPIVLPFLMGIMDASEVCIIEYFKDTNILFVGLGMVAIAVEHNLHKRIALQVLLIIGVR 120
Qy 121 PALLLGFMLVTAFLSMWISNTATTAMVPIGHAVLEQLOQSKKDVGGNNPTFELOEE 178
Db 121 PALLLGFMLVTAFLSMWISNTATTAMVPIGHAVLEQLOQSKKDVGGNNPTFELOEE 178
Qy 179 EECPOKEVTKLDNGQ--PV-SAPSEPRTKTQEHHRFSQGLSLCICYSASIGGIATLTGT 235
Db 179 EECPOKEVTKLDNGQ--PV-SAPSEPRTKTQEHHRFSQGLSLCICYSASIGGIATLTGT 235
Qy 191 EPSQKEVTKLDNGQALPVTSSSEGRAHLQSKGHUHLTQCMSCVCSIGGIATLTGT 240
Db 191 EPSQKEVTKLDNGQALPVTSSSEGRAHLQSKGHUHLTQCMSCVCSIGGIATLTGT 240
Qy 236 TPNLVLOQVNSLPQNGNVNPFASWFGFAPPTMIILLALLAWLQVLFLGVNFRKNFGF 295
Db 241 APNLVLOQVNSLPQNGNVNPFASWFGFAPPTMIILLALLAWLQVLFLGVNFRKNFGF 300
Qy 296 GEGEERKQAAFOVIKQYRLGPMSPFAEKTVTVLVLLVVLVLTREPFGFGMDTVFANEKQ 355
Db 301 GEGEERKQAAFOVIKQYRLGPMSPFAEKTVTVLVLLVVLVLTREPFGFGMDTVFANEKQ 360
Qy 356 NEKGOSMASDGTVAIFSLVMFIIPSKIPGLMQDPKPKGLKAPPALLTWKTVNDKMPN 415
Db 361 NAGESMWSVSDGTVAIFGIIMFIIPSKIPGLMQDPKPKGLKAPPALLTWKTVNDKMPN 420
Qy 416 IVILLGGGFALAKGSEGLSEWLGDKLTPLQHIPPPSATAVILCLLIAIFTECTSNVATT 475
Db 421 IVILLGGGFALAKGSEGLSEWLGDKLTPLQHIPPPSATAVILCLLIAIFTECTSNVATT 480
Qy 476 TFLPILASMAQAICLHPLYVMLPCTLAASLAFMLPVATPPNAIVFSGGLKVSMDMARAG 535
Db 481 TFLPILASMAQAICLHPLYVMLPCTLAASLAFMLPVATPPNAIVFSGGLKVSMDMARAG 540
Qy 536 FLNIIIGVLAITLSINSWSIPFKLDTFPPSWAHNS--TSQCLLNPSNSTV 584
Db 541 FLNIIIGVLAITLSINSWSIPFKLDTFPPSWAHNS--TSQCLLNPSNSTV 590

```

RESULT 5
US-10-167-994-11
; Sequence 11, Application US/10167994
; Publication No. US20030082647A1
; GENERAL INFORMATION:
; APPLICANT: Reenan, Robert A.
; APPLICANT: Rogina, Blanka
; APPLICANT: Helfand, Stephen L.
; TITLE OF INVENTION: TRANSPORTER PROTEIN
; FILE REFERENCE: 13407-013001
; CURRENT APPLICATION NUMBER: US/10/167,994
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/255,013
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 10/017,479
; PRIOR FILING DATE: 2001-12-12
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 592
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-167-994-11

Query Match      78.2%; Score 2392; DB 14; Length 592;
Best Local Similarity 76.9%; Pred. No. 2e-217;
Matches 454; Conservative 63; Mismatches 67; Indels 6; Gaps 4;

QY 1 MATCPALWAYRFLIIVLCPIFLPLIVOTKEAYCAYSIILMALLMCTEALPLAVTA 60
DB 1 MATCQALWAYRSLLIVFFPILLPLILVPSKEAYCAVAIILMALFWCTEALPLAVTA 60

QY 61 LPFIVLPALMIDASEVCIEYFKDNLIFVGGMLVAIAVEHNLHRIALQVLLIGVR 120
DB 61 LPFILFPFMGIVDAVEAYEYLKDSNLLFFGGLLVAIAVEHNLHRIALRVLLIGVR 120

QY 121 PALILGFMVLAFLSMWISNTATTAMVPVIGHAVLEQLOSK--KDVGGNNPTFELQ 178
DB 121 PAPILGFMVLAFLSMWISNTATTAMVPVIGHAVLEQLOSK--KDVGGNNPTFELQ 180

QY 179 BECPQKEVTKLDNGQ--PV-SAPSEPRTKTQEHHRFSQGLSLCICYSASIGGIATLTGT 235
DB 181 EPSQKEVTKLDNGQALPVTISASSEGRAHLSQKHLHLTQCWSLVCVYSASIGGIATLTGT 240

QY 236 TPNVLQGVNSLFPQNGNVNPFASWFGFAPPTMIILLIALLWLVLFVGNFRKNFGF 295
DB 241 APNLVLOQVNSLFPQNGNVNPFASWFGFAPPTMIILLIALLWLVLFVGNFRKNFGI 300

QY 296 GEGEERKQAAFOVTKTOYRLLGPMSPFAEKTIVTVLVLLVLMFTREPGFPFGMDTVFA 355
DB 301 GEKQBEQQAAYCVIQTREHRLGPMTPFAEKASISILFVILVLLWFTREPGFFLGWNLAFP 360

QY 356 NEKQSMASDGTVAIFISLVNFIIPSKIPLGMQDPKPGKIKAPPAILLTWKTVNDKMPWN 415
DB 361 NAKGESMVSDGTVAIFIGIIMFIIPSKPPGLTQDPENPGKIKAPGLDLDKTVNQKMPWN 420

QY 416 IVILLGGFALAKSEQSGLSWLDKLTPLQHTPPSNATAVILCLLTAIETECTSNVATT 475
DB 421 IVLLGGGYALAKSERSGLSEWLGKLTPLQSPVAPAIILSLLVATFTECTSNVATT 480

QY 476 TLEPILASMAQAICLHPYVYMLPCTLAASLAFMLPVATPPNAIVFSGGLKVSMDMARAG 535
DB 481 TIFPILASMAQAICLHPYVYMLPCTLATSLAFMLPVATPPNAIVFSGGLKVLDMARAG 540

QY 536 FLLNIIGVLAITLSINSWSIPKLDTPPSWAHNS-TSQCLLNPSNSTVP 584
DB 541 FLLNIIGVLAITLSINSWSIPKLDTPPSWAHNS-TSQCLLNPSNSTVP 590

RESULT 6
US-09-729-094-4
; Sequence 4, Application US/09729094
; GENERAL INFORMATION:
; APPLICANT: CHATURVEDI, Kabir et al
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; FILE REFERENCE: CL000662
; CURRENT APPLICATION NUMBER: US/09/729,094
; CURRENT FILING DATE: 2000-12-05
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 619
; TYPE: PRT
; ORGANISM: Xenopus laevis
US-09-729-094-4

Query Match      64.5%; Score 1971.5; DB 9; Length 619;
Best Local Similarity 63.2%; Pred. No. 1.6e-177;
Matches 386; Conservative 87; Mismatches 85; Indels 53; Gaps 10;

QY 10 AYRFYLVLCPLIFLPLPLIVOTKEAYCAYSIILMALLMCTEALPLAVTALFPIVLPFL 69
DB 10 ANRNYFIIPLVPLFLPLVVPVPTKEASCGFVIVWALFWCTEALPLAVTALFPIVLPFL 69

QY 70 MGIMDASEVCIEYFKDNLIFVGGMLVAIAVEHNLHRIALQVLLIGVRPALLLGM 129
DB 70 MGIMDSTAVCSQYLDKTNMLFVGGMLVAIAVEHNLHRIALRVLLIGVVKPALLLGM 129

QY 130 LVTAFLSMWISNTATTAMVPVIGHAVLEQLOSK-----KDVGGNN----- 171
DB 130 VVTAFLSMWISNTATTAMVPVIGHAVLEQLOSK-----KDVGGNN----- 189

QY 172 -----NPTFELQECC-----POKEVTKLDNGQPVSAAP-SEPRTKT 206
DB 190 SVNPSGKMLAINDTVATENEFEIQEKTDPESKQ-EKOSIGIVIEPEDEKQTEEK 248

QY 207 Q--EHRFSGQLSLCICYSASIGGIATLTGTPNLVLOQVNSLFPQNGNVNPFASWFGF 264
DB 249 QKEKHLKICKGMSLVCVYSASIGGIATLTGTPNLVLOQVNSLFPQNGNVNPFASWFGF 308

QY 265 APTMIILLIALLWLVLFVGNFRKNFGFEGEER--ERKQAAFOVTKTOYRLLGPMSP 322
DB 309 APTMLVLLALSMLWLFQTYLVGNFKNFGCGGNAEQEKEKRAFRVISGEHKKLGSMTF 368

QY 323 AEKTVTVLVLLVLMFTREPGFPFGMDTVFAVEHNLHRIALQVLLIGVRPALLLGM 382
DB 369 AEISVLVILLVLLWFTREPGFPGMA-TISFNKGGKEMVTDATVAIFVSLMMFFPSE 427

QY 383 IPGL-MQDPKPG---KLKAPPAILLTWKTVNDKMPWNIVLLGGGFALAKSEQSGLSW 438
DB 428 LPSFKYQDTPKPGKPLRVPPALLDWKTVNERMPWNIVLLGGGFALAKSEQSGLSW 487

QY 439 LGDKLTPLQHTPPSNATAVILCLLTAIETECTSNVATTTLFPLASMAQAICLHPYVML 498
DB 488 LGSEKLTPLQHTPPSNATAVILCLLTAIETECTSNVATTTLFPLASMAQAICLHPYVML 547

QY 499 PCTLAASLAFMLPVATPPNAIVFSGGLKVSMDMARAGFLNIIGVLAITLSINSWSIP 558
DB 548 PCTLSLAFMLPVATPPNAIVFSGGLKVSMDMARAGFLNIIGVLAITLSINSWSIP 607

QY 559 KLDTPPSWAHS 569
DB 608 NLGTFPSWANA 618

RESULT 7
US-10-435-631-4
; Sequence 4, Application US/10435631
; Publication No. US20030186381A1
; GENERAL INFORMATION:
; APPLICANT: CHATURVEDI, Kabir et al
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,

```

```

RESULT 8
US-10-718-359-12
; Sequence 12, Application US/10718359
; Publication No. US20050095240A1
; GENERAL INFORMATION:
; APPLICANT: MEDICAL COLLEGE OF GEORGIA RESEARCH INSTITUTE, INC.
; TITLE OF INVENTION: NACT AS A TARGET FOR LIFESPAN EXPANSION AND WEIGHT REDUCTION
; FILE REFERENCE: 275.0008 0101
; CURRENT APPLICATION NUMBER: US/10/718,359
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/428,469

```

RESULT 9

US-10-092-900A-46
; Sequence 46, Application US/10092900A
; Publication No. US20040043382A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Taupier Jr., Raymond J.
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Li, Li
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Ji, Weizhen
; APPLICANT: Gorman, Linda
; APPLICANT: Miller, Charles E.
; APPLICANT: Kekuda, Ramsekh

QY	DB	APPLICANT:	SEQUENCE	SCORE	DB 15	LENGTH	616	7
241	QCQEDQERKELCKAMTLCICVAAISIGTATLTGTGPNVLLGQNNELFPDSKDLVNFASW	241	QCQEDQERKELCKAMTLCICVAAISIGTATLTGTGPNVLLGQNNELFPDSKDLVNFASW	300				
262	FGFAPPTMIILLLAWLWQVLFGVNNFRNFKFGEGEERKQAAFOVITQYRLLGPM	262	FGFAPPTMIILLLAWLWQVLFGVNNFRNFKFGEGEERKQAAFOVITQYRLLGPM	321				
301	FAFAFPNNMLVLLFAWLQVFMVYMFSSFKKSGWGLSKKNEKAALVKVQBEYKLGPLS	301	FAFAFPNNMLVLLFAWLQVFMVYMFSSFKKSGWGLSKKNEKAALVKVQBEYKLGPLS	360				
322	FAEKTVTVLVLLVLMFTREPFGPCWGDITVPANFKGQSMASDGTVAIFSLVMFIIPS	322	FAEKTVTVLVLLVLMFTREPFGPCWGDITVPANFKGQSMASDGTVAIFSLVMFIIPS	381				
361	FAEINVLCIFLLVILWFSRDPGFMPOGLTVAVVEGTKSV-SDATVAIFVATLLFVIPS	361	FAEINVLCIFLLVILWFSRDPGFMPOGLTVAVVEGTKSV-SDATVAIFVATLLFVIPS	419				
382	KIP-----GLMQDPKPKGLKAPPAITLTKVTNDKMPNIVILLGGGFALAKGSEQSGLS	382	KIP-----GLMQDPKPKGLKAPPAITLTKVTNDKMPNIVILLGGGFALAKGSEQSGLS	436				
420	QKPRFNFRSQTECKSP-VLIAPPPLLDWKVQSKVPWGI VLLGGGFALAKGSEASGLS	420	QKPRFNFRSQTECKSP-VLIAPPPLLDWKVQSKVPWGI VLLGGGFALAKGSEASGLS	478				
437	EWLGDKLTPLQHIIPPSATAVILCLLIAIFTECTSNVATTLFLPILASMAQAICLHPLVY	437	EWLGDKLTPLQHIIPPSATAVILCLLIAIFTECTSNVATTLFLPILASMAQAICLHPLVY	496				
479	VVMGKQMEPLHAVPPAAITLILSLVAVFTECTSNVATTLFLPILASMAQAICLHPLVY	479	VVMGKQMEPLHAVPPAAITLILSLVAVFTECTSNVATTLFLPILASMAQAICLHPLVY	538				
497	MLPCTLAASLAFMLPVATPPNAIVFSPGGLKVKVSDMARAGFLNIIIGVLAITLSINSWIP	497	MLPCTLAASLAFMLPVATPPNAIVFSPGGLKVKVSDMARAGFLNIIIGVLAITLSINSWIP	556				
539	MLPCTLSASFAPMLPVATPPNAIVFTYGLKVKVADWMTGVTIMNIIIGVFCVFLAVNTWGRA	539	MLPCTLSASFAPMLPVATPPNAIVFTYGLKVKVADWMTGVTIMNIIIGVFCVFLAVNTWGRA	598				
557	IFKLDTPPSMAH 568	557	IFKLDTPPSMAH 568					
599	IFDLDFPDWAN 610	599	IFDLDFPDWAN 610					
RESULT 10								
US-10-092-900A-48								
; Sequence 48, Application US/10092900A								
; Publication No. US20040043382A1								
; GENERAL INFORMATION:								
; APPLICANT: Padigaru, Muralidhara								
; APPLICANT: Spytek, Kimberly A.								
; APPLICANT: Shenoy, Suresh G.								
; APPLICANT: Taupier Jr., Raymond J.								
; APPLICANT: Pena, Carol E.A.								
; APPLICANT: Li, Li								
; APPLICANT: Zerhusen, Bryan D.								
; APPLICANT: Gusev, Vladimir Y.								
; APPLICANT: Ji, Weizhen								
; APPLICANT: Gorman, Linda								
; APPLICANT: Miller, Charles E.								
; APPLICANT: Kekuda, Ramesh								
; APPLICANT: Patturajan, Meera								
; APPLICANT: Gangolli, Esha A.								
; APPLICANT: Vernet, Corine A.M.								
; APPLICANT: Guo, Xiaojia Sasha								
; APPLICANT: Tchernev,, Velizar T.								
; APPLICANT: Fernandes, Elma R.								
; APPLICANT: Casman, Stacie J.								
; APPLICANT: Malyankar, Uriel M.								
; APPLICANT: Gerlach, Valerie								
; APPLIC								

```
/ PRIOR APPLICATION NUMBER: USSN 60/274,281
/ PRIOR FILING DATE: 2001-03-08
/ PRIOR APPLICATION NUMBER: USSN 60/274,191
/ PRIOR FILING DATE: 2001-03-08
/ PRIOR APPLICATION NUMBER: USSN 60/325,681
/ PRIOR FILING DATE: 2001-09-27
/ PRIOR APPLICATION NUMBER: USSN 60/304,354
/ PRIOR FILING DATE: 2001-07-10
/ PRIOR APPLICATION NUMBER: USSN 60/279,995
/ PRIOR FILING DATE: 2001-03-30
/ PRIOR APPLICATION NUMBER: USSN 60/294,899
/ PRIOR FILING DATE: 2001-05-31
/ PRIOR APPLICATION NUMBER: USSN 60/287,424
/ PRIOR FILING DATE: 2001-04-30
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 768
/ SEQ ID NO 48
/ LENGTH: 616
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-092-900A-48

Query Match          53.2%; Score 1628; DB 15; Length 616;
Best Local Similarity 51.3%; Pred. No. 5.7e-145;
Matches 314; Conservative 113; Mismatches 139; Indels 46; Gaps 7;

QY 1 MATCHPALWAYRYLVLCPLIFLLPLIVOTKEAYCAYSIILMALLMCTEALPLAVTA 60
  || : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 MASALSYVSKPKSVFLVFTPLLLPLVILMPAKVSCCAVYIILMAIYWCTEVIPLAVTS 60

QY 61 LPPIVLPMLGMDASEVCIEYFKDTNLLFVGLGMVAIAVEHNLHKKRIALQVLLIIGVR 120
  || : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 LMPVLLPFLQILDSRQVCVQYMKDNTNMLFGLGLIVAVAVERNLHKKRIALTLLWVGAK 120

QY 121 PALLLGLFMLVTAFLSMWISNTATTAMVPIGHAVLEOLQGSKKDVEGG-----NNN 172
  || : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 121 PARMLGFMGVTAFLSMWISNTATTAMVPIVEAILQMEATSAATEAGLEGQGTINNL 180

QY 173 PTFE-----LQBECPQ-----KEVTKL-----DNGQPVSAPEP 201
  || : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 181 NALEDDETVAIVGGKCVAILIISTYVKKVEKLQINNLMTPKKLEKEQEQDLPGRIPQDSA 240

QY 202 RTQKTQEHHRFSQGLSLCICYSASIGGIATLTGTTNPLVLOGQVNSLPQNGNVNPFASW 261
  || : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 241 QCQEDQERKRLCKAMTLCICYAASIGGTATLTGTGPNVLLQGMNELFPDSKDLVNFASW 300

QY 262 FGPAFTMIILLIALLWMLQVLFVGNFRKNFGFGEGERKQAAFOVKTQYRLLGPM 321
  || : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 301 FAFAPFNMLVLLFAWLQVLYMFSSFKSGWGLSKKNEKAAALKVLOEYRKLGLPLS 360

QY 322 KIP-----GLMQDPKKPKGLKAPPAILTWTNDKMPNIVILLGGGFALAKGSEQSGLS 436
  || : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 420 QKPKFNFRSQTSEEGKSP-VLIAPPPLLDWKVQKVPWGVLLGGGFALAKGSEQSGLS 478

QY 437 EWLGDKLTPLOHIPSATAVILCLLIAIPTCTSNVATTTPLPLASMAOAILCHPIYV 496
  || : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 479 VMGQKMEPLHAPVPAATLTLSSLVAVFTCTSNVATTTPLPLFASMSRSISGLNPLYI 538

QY 497 MLPCTLAASLAFMLPVATPPNAIVPSFGGLKVSMDARAGFLNIIIGVLAITLSINSWSIP 556
  || : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 539 MLPCTLSASFAMLPVATPPNAIVFTYGHKLKADVMKTVGMNIIIGVFCVFLAVNTWGRA 598

QY 557 IFKLDTPPSWAH 568
  || : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 599 IFDLDPDWMAN 610
```

RESULT 11
US-10-403-161-68

```
/ Sequence 68, Application US/10403161
/ Publication No. US20040043930A1
/ GENERAL INFORMATION:
/ APPLICANT: Anderson, David et al.
/ TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
/ FILE REFERENCE: 21402-573C
/ CURRENT APPLICATION NUMBER: US/10/403,161
/ CURRENT FILING DATE: 2003-03-31
/ PRIOR APPLICATION NUMBER: 60/370349
/ PRIOR FILING DATE: 2002-04-05
/ PRIOR APPLICATION NUMBER: 60/384543
/ PRIOR FILING DATE: 2002-05-30
/ PRIOR APPLICATION NUMBER: 60/370969
/ PRIOR FILING DATE: 2002-04-08
/ PRIOR APPLICATION NUMBER: 60/403748
/ PRIOR FILING DATE: 2002-08-15
/ PRIOR APPLICATION NUMBER: 60/372019
/ PRIOR FILING DATE: 2002-04-12
/ PRIOR APPLICATION NUMBER: 60/374379
/ PRIOR FILING DATE: 2002-04-22
/ PRIOR APPLICATION NUMBER: 09/779679
/ PRIOR FILING DATE: 2001-02-08
/ PRIOR APPLICATION NUMBER: 60/181045
/ PRIOR FILING DATE: 2000-02-08
/ PRIOR APPLICATION NUMBER: 10/055877
/ PRIOR FILING DATE: 2002-01-22
/ PRIOR APPLICATION NUMBER: 60/262892
/ PRIOR FILING DATE: 2001-01-19
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 173
/ SOFTWARE: Curaseqlist version 0.1
/ SEQ ID NO 68
/ LENGTH: 616
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-403-161-68

Query Match          53.2%; Score 1628; DB 15; Length 616;
Best Local Similarity 51.3%; Pred. No. 5.7e-145;
Matches 314; Conservative 113; Mismatches 139; Indels 46; Gaps 7;

QY 1 MATCHPALWAYRYLVLCPLIFLLPLIVOTKEAYCAYSIILMALLMCTEALPLAVTA 60
  || : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 MASALSYVSKPKSVFLVFTPLLLPLVILMPAKVSCCAVYIILMAIYWCTEVIPLAVTS 60

QY 61 LPPIVLPMLGMDASEVCIEYFKDTNLLFVGLGMVAIAVEHNLHKKRIALQVLLIIGVR 120
  || : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 LMPVLLPFLQILDSRQVCVQYMKDNTNMLFGLGLIVAVAVERNLHKKRIALTLLWVGAK 120

QY 121 PALLLGLFMLVTAFLSMWISNTATTAMVPIGHAVLEOLQGSKKDVEGG-----NNN 172
  || : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 121 PARMLGFMGVTAFLSMWISNTATTAMVPIVEAILQMEATSAATEAGLEGQGTINNL 180

QY 173 PTFE-----LQBECPQ-----KEVTKL-----DNGQPVSAPEP 201
  || : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 181 NALEDDETVAIVGGKCVAILIISTYVKKVEKLQINNLMTPKKLEKEQEQDLPGRIPQDSA 240

QY 202 RTQKTQEHHRFSQGLSLCICYSASIGGIATLTGTTNPLVLOGQVNSLPQNGNVNPFASW 261
  || : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 241 QCQEDQERKRLCKAMTLCICYAASIGGTATLTGTGPNVLLQGMNELFPDSKDLVNFASW 300

QY 262 FGPAFTMIILLIALLWMLQVLFVGNFRKNFGFGEGERKQAAFOVKTQYRLLGPM 321
  || : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 301 FAFAPFNMLVLLFAWLQVLYMFSSFKSGWGLSKKNEKAAALKVLOEYRKLGLPLS 360

QY 322 PAEKTVTVLVLLVLTREPGFPFGMDTVFANEKQSGMASDGTVAIFISLVNFIIPS 381
  || : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 361 FAENVLICFFLLVILWFSRDPGFMPCWLTVAWVEGETKSV-SDATVAIFVATLLFVIPS 419

QY 382 KIP-----GLMQDPKKPKGLKAPPAILTWTNDKMPNIVILLGGGFALAKGSEQSGLS 436
  || : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 420 QKPKFNFRSQTSEEGKSP-VLIAPPPLLDWKVQKVPWGVLLGGGFALAKGSEQSGLS 478
```

QY 437 EWLGDKLTPLOHIPPSSATAVILLCIIAIFTECTSNVATTFLPILASMAQAICLHPLVY 496
 Db 479 VMGKQMEPLHVPAPPAITLISLVAVTECTSNVATTFLPILASMSRSIGLNPVLI 538
 QY 497 MLPCTLLAASLAFMLPVATPPNAIVFSGGLKVDMARAGFLNIIGVLAITLINSWSIP 556
 Db 539 MLPCTLLSASFAFMLPVATPPNAIVFTYGHKLVADVMKTVGIMNIIGVFCVFLAVNTWGRA 598
 QY 557 IFKLDTFPPSWAH 568
 Db 599 IFDLDFPDWAN 610

RESULT 12
 US-10-173-519-2
 ; Sequence 2, Application US/10173519
 ; Publication No. US20020193582A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Curtis, Rory A.J.
 ; TITLE OF INVENTION: 59624, A Human Transporter Family Member
 ; TITLE OF INVENTION: and Uses Therefor
 ; FILE REFERENCE: MP101-098P1RM
 ; CURRENT APPLICATION NUMBER: US/10173,519
 ; PRIOR FILING DATE: 2002-06-17
 ; PRIOR APPLICATION NUMBER: 60/298,970
 ; PRIOR FILING DATE: 2001-06-18
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 568
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-173-519-2

Query Match 53.1%; Score 1624; DB 13; Length 568;
 Best Local Similarity 53.2%; Pred. No. 1.2e-144;
 Matches 309; Conservative 108; Mismatches 132; Indels 32; Gaps 6;

QY 1 MATCWPALWAYRFLVLCIPFLPLPLIVOTKEAYCAYSIIMALLWCTEALPLAVTA 60
 Db 1 MASALSYSVKFSFVLFTVPLLLPLVILMPAKFVRCAYVILMAIYCTEVIPLAVTS 60
 QY 61 LPFIVLPFLMGINDASEVCIEYKDNILFVGGMLVAIAVEHNLKRIALQVLLIIGVR 120
 Db 61 LMPVLLPFLQIDLSRQVCQVMKDNMLFLGLLIIVAVAVERNLHKRIALRIALLWVGAK 120
 QY 121 PALLLGFMLVTAFLSMWISNTATTAMVPPIGHAVLEQLOGSKKDVGGNNPTFELQEE 180
 Db 121 PARMLGFMGVMTALLSMWISNTATTAMVPPIVEAILQOMEATSAATEAG-----LELVDK 175
 QY 181 CPQKEVTKLNDGQPVASPSRPTQKQEHHRFSGGLSLCICYSASIGGIATLTGTPNLV 240
 Db 176 GKAKE---LPGSQVIFEGPTLGGQEDQERKLCAMTKLCICYAASIGGTATLTGTGNV 232
 QY 241 LQGVNSLFPQNGVNVNFASFWEFAFPTMILLLAWMLQVLPGLVNFKNFGFGE 300
 Db 233 LLQGMNELFDSKDLNVNFASFWEFAFPMNLVMLLFAWMLQFYVMRFNFKKSGCGLESK 292
 QY 301 ERKQAAFOVKTQVRLGPMSPFAEKTVTVLVLLVWLFTRPFGPGW-----GDTV 353
 Db 293 KNEKAALKVQEEYRKLGLPSFAEINVLICFFLLVILWFSRDPGFMPLTVAWVEGETK 352

Db 462 CTSNVATTTLFLPILFASMSRSIGLNPVILMPLCTLLSASFAFMLPVATPPNAIVFTYGHULK 521
 QY 528 VSDMARAGFLNIIGVLAITLINSWSIPIFKLDTFPPSWAH 568
 Db 522 VADMVKTVGIMNIIGVFCVFLAVNTWGRAIFDLDFPDWAN 562

RESULT 13
 US-10-403-161-72
 ; Sequence 72, Application US/10403161
 ; Publication No. US20040043950A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Anderson, David et al.
 ; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
 ; FILE REFERENCE: 21402-573C
 ; CURRENT APPLICATION NUMBER: US/10/403,161
 ; CURRENT FILING DATE: 2003-03-31
 ; PRIOR APPLICATION NUMBER: 60/370349
 ; PRIOR FILING DATE: 2002-04-05
 ; PRIOR APPLICATION NUMBER: 60/384543
 ; PRIOR FILING DATE: 2002-05-30
 ; PRIOR APPLICATION NUMBER: 60/370969
 ; PRIOR FILING DATE: 2002-04-08
 ; PRIOR APPLICATION NUMBER: 60/403748
 ; PRIOR FILING DATE: 2002-08-15
 ; PRIOR APPLICATION NUMBER: 60/372019
 ; PRIOR FILING DATE: 2002-04-12
 ; PRIOR APPLICATION NUMBER: 60/374379
 ; PRIOR FILING DATE: 2002-04-22
 ; PRIOR APPLICATION NUMBER: 09/779679
 ; PRIOR FILING DATE: 2001-02-08
 ; PRIOR APPLICATION NUMBER: 60/181045
 ; PRIOR FILING DATE: 2000-02-08
 ; PRIOR APPLICATION NUMBER: 10/055877
 ; PRIOR FILING DATE: 2002-01-22
 ; PRIOR APPLICATION NUMBER: 60/262892
 ; PRIOR FILING DATE: 2001-01-19
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 173
 ; SOFTWARE: CuraseqList version 0.1
 ; SEQ ID NO 72
 ; LENGTH: 568
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-403-161-72

Query Match 53.1%; Score 1624; DB 15; Length 568;
 Best Local Similarity 53.2%; Pred. No. 1.2e-144;
 Matches 309; Conservative 108; Mismatches 132; Indels 32; Gaps 6;

QY 1 MATCWPALWAYRFLVLCIPFLPLPLIVOTKEAYCAYSIIMALLWCTEALPLAVTA 60
 Db 1 MASALSYSVKFSFVLFTVPLLLPLVILMPAKFVRCAYVILMAIYCTEVIPLAVTS 60
 QY 61 LPFIVLPFLMGINDASEVCIEYKDNILFVGGMLVAIAVEHNLKRIALQVLLIIGVR 120
 Db 61 LMPVLLPFLQIDLSRQVCQVMKDNMLFLGLLIIVAVAVERNLHKRIALRIALLWVGAK 120
 QY 121 PALLLGFMLVTAFLSMWISNTATTAMVPPIGHAVLEQLOGSKKDVGGNNPTFELQEE 180
 Db 121 PARMLGFMGVMTALLSMWISNTATTAMVPPIVEAILQOMEATSAATEAG-----LELVDK 175
 QY 181 CPQKEVTKLNDGQPVASPSRPTQKQEHHRFSGGLSLCICYSASIGGIATLTGTPNLV 240
 Db 176 GKAKE---LPGSQVIFEGPTLGGQEDQERKLCAMTKLCICYAASIGGTATLTGTGNV 232
 QY 241 LQGVNSLFPQNGVNVNFASFWEFAFPTMILLLAWMLQVLPGLVNFKNFGFGE 300
 Db 233 LLQGMNELFDSKDLNVNFASFWEFAFPMNLVMLLFAWMLQFYVMRFNFKKSGCGLESK 292
 QY 301 ERKQAAFOVKTQVRLGPMSPFAEKTVTVLVLLVWLFTRPFGPGW-----GDTV 353
 Db 293 KNEKAALKVQEEYRKLGLPSFAEINVLICFFLLVILWFSRDPGFMPLTVAWVEGETK 352

Db 462 CTSNVATTTLFLPIFASMSRSIGLNPLYIMLPCTLSASFAFMLPVATPPNAIVFTYGHK 521

Qy 528 VSDMARAGFLNIIGVLAITLSINSWSPIFKLDTPPSWAH 568

Db 522 VADMVKTGVIMNIIGVFCVFLAVNTWGRAIFDLDFPDWAN 562

Search completed: June 30, 2005, 09:32:50
Job time : 78.171 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2005, 08:29:45 ; Search time 17.5314 Seconds
(without alignments)
3221.603 Million cell updates/sec

Title: US-10-017-479A-4
Perfect score: 3058
Sequence: 1 MATCHPALWAYRYLVILCL.....HSNTSQCLLNPSNVTVPGL 587

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : PIR_79: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2340.5	76.5	593	2 J46528	sodium/dicarboxyla
2	1622	53.0	568	2 I47911	Na+-coupled citrat
3	1259	41.2	595	2 A47714	Na+/sulfate cotran
4	1037.5	33.9	526	2 S43561	YCR37C homolog K08
5	999.5	33.9	534	2 G88575	protein K08E5.2 [i
6	999.5	32.7	599	2 T21613	hypothetical prote
7	927	30.3	545	2 T18694	hypothetical prote
8	907.5	29.7	520	2 C89980	hypothetical prote
9	830.5	27.2	539	2 A88546	protein R107.1 [im
10	814.5	26.6	539	2 S30871	hypothetical prote
11	785	25.7	548	2 E71961	probable transport
12	781	25.5	552	2 F64546	sodium-dependent t
13	556	18.2	461	2 I64080	probable membrane
14	548	17.9	446	2 F70302	transporter (Pho87
15	525.5	17.2	462	2 B22510	transporter, Nadc
16	500.5	16.4	471	2 G81156	transporter, Nadc
17	496.5	16.2	456	2 AF0093	Sodium, sulfate sym
18	494.5	16.2	471	2 A81948	probable transmem
19	478.5	15.6	443	2 C69205	sodium/dicarboxyla
20	468.5	15.3	487	2 C82215	probable transport
21	457	14.9	432	2 H64383	Na+ transporter -
22	455	14.9	478	2 A82336	probable transport
23	441	14.4	157	2 S36784	mucin - rat (fragm
24	383.5	12.5	517	2 C99840	conserved hypothet
25	344	11.2	867	2 T40336	probable MSF trans
26	313	10.2	881	2 S46633	probable membrane
27	305	10.0	513	2 I39534	hypothetical prote
28	305	10.0	894	2 S45135	probable membrane
29	294	9.6	612	2 S74936	sulfur deprivation

30	290	9.5	450	2 A10909	probable membrane
31	288	9.4	923	1 MMBY7C	probable membrane
32	256.5	8.4	487	1 B64795	ybdS protein - Esc
33	256.5	8.4	487	2 G85560	probable membrane
34	256.5	8.4	487	2 C90710	probable membrane
35	242.5	7.9	589	2 H84072	sodium/sulfate sym
36	240.5	7.9	501	2 AE0578	citrate carrier [i
37	235.5	7.7	487	1 QOECRS	yglJ protein - Esc
38	234.5	7.7	487	2 B91122	hypothetical prote
39	234.5	7.7	487	2 A85967	hypothetical prote
40	228	7.5	424	2 F71205	hypothetical prote
41	223.5	7.3	608	2 AC0798	probable sodium/su
42	221	7.2	610	2 A85870	probable transport
43	221	7.2	610	2 H91025	probable transport
44	221	7.2	610	2 B65001	probable transport
45	220.5	7.2	610	2 AG0312	probable ion trans
46	213	7.0	425	2 C75019	transport protein
47	200.5	6.6	488	2 A81194	C4-dicarboxylate t
48	200.5	6.6	488	2 B81831	probable integral
49	200.5	6.6	610	2 H83165	probable sodium/su
50	197	6.4	618	2 F84409	arsenite transport
51	196.5	6.4	589	2 F75398	transporter, sodiu
52	188.5	6.2	484	2 AE0887	probable membrane
53	185	6.0	479	1 B64043	conserved hypothet
54	177.5	5.8	477	1 B64813	ybh1 protein - Esc
55	177.5	5.8	590	2 AG3106	transporter Atu447
56	177.5	5.8	590	2 E98180	transporter, sodiu
57	173.5	5.7	477	2 G85579	probable membrane
58	173.5	5.7	477	2 F90728	probable membrane
59	170.5	5.6	592	2 D95393	hypothetical prote
60	164.5	5.4	471	2 F71543	probable dicarboxy
61	160.5	5.2	476	2 B97096	2-oxoglutarate/mal
62	160	5.2	470	2 G86516	dicarboxylase tran
63	159	5.2	469	2 T35526	probable integral
64	157.5	5.2	482	2 F71969	hypothetical prote
65	157	5.1	472	2 F90078	hypothetical prote
66	155.5	5.1	585	2 S69216	sulfur deprivation
67	154	5.0	470	2 H72105	dicarboxylase tran
68	153	5.0	838	2 S28911	gene DN10 protein
69	148.5	4.9	475	2 A10010	probable membrane
70	147.5	4.8	411	2 G90154	arsenite transport
71	143	4.7	478	1 F69811	2-oxoglutarate/mal
72	142	4.6	436	2 S47723	arsenite efflux pu
73	139	4.5	429	2 F91175	arsenical pump mem
74	139	4.5	429	2 F86021	arsenical pump mem
75	135	4.4	513	2 A99839	Na+/H+-exchanging
76	135	4.4	513	2 G84884	Na+/H+-exchanging
77	134.5	4.4	425	2 F72315	conserved hypothet
78	134	4.4	513	2 A85697	Na+/H+ antiporter,
79	132.5	4.3	431	2 A11599	arsenic efflux pum
80	131.5	4.3	532	2 A57173	oculocutaneous alb
81	129	4.2	603	2 B59154	NADH2 dehydrogenas
82	129	4.2	603	2 T14026	NADH2 dehydrogenas
83	128.5	4.2	429	1 C41903	arsenical pump mem
84	128	4.2	515	2 E72089	ADP, ATP carrier p
85	126.5	4.1	429	1 B41902	arsenical pump mem
86	126.5	4.1	522	2 B64067	Na+/H+-exchanging
87	126	4.1	461	2 AB2185	hypothetical prote
88	125	4.1	430	2 G89982	hypothetical prote
89	125	4.1	515	2 B66534	ADP/ATP translocas
90	125	4.1	620	2 H82761	sulfur deprivation
91	124	4.1	614	2 AG0310	NADH2 dehydrogenas
92	123.5	4.0	548	2 AC3587	sulfur deprivation
93	122	4.0	429	2 G70528	probable arsa prot
94	122	4.0	514	2 AG0723	regulator of intra
95	120.5	3.9	429	1 B25937	arsenical pump mem
96	118	3.9	603	1 B2HUN5	NADH2 dehydrogenas
97	116.5	3.8	370	2 AB0602	probable membrane
98	116.5	3.8	401	2 AF1800	transport protein
99	116	3.8	417	2 AF2992	arsenical pump mem
100	116	3.8	417	2 D98291	probable arsenical

ALIGNMENTS

RESULT 1
146528
sodium/dicarboxylate cotransporter - rabbit (fragment)
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 14-Feb-1997 #sequence_revision 14-Feb-1997 #text_change 09-Jul-2004
C:Accession: 146528
R:Pajor, A.M.
J. Biol. Chem. 270, 5779-5785, 1995
A:Title: Sequence and functional characterization of a renal sodium/dicarboxylate cotransporter
A:Reference number: 146528; MUID:95197598; PMID:7890707
A:Accession: 146528
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-593 <PAB>
A:Cross-references: UNIPROT:Q28615; EMBL:U12186; NID:g758383; PID:G758383
C:Superfamily: sodium/sulfate cotransporter

Query Match 76.5%; Score 2340.5; DB 2; Length 593;
Best Local Similarity 74.0%; Pred. No. 1.6e-181;
Matches 439; Conservative 71; Mismatches 74; Indels 9; Gaps 5;
QY 1 MATCWPAALWAYRFLYIVLCPLPILPLIVOTKEAYCAYSIIMALLWCTEALPLAVTA 60
DB 1 MATCWQGLWAYRMVLLVPLLSISLLPLILVPRKEAYCAVAIIMLWFWCTDALPLAVTA 60
QY 61 LPPIVLPLMGIMDASEVCIEYFKDTNLLFVGLMVAIAVEHNLHKLRIALQVLLIIGVR 120
DB 61 LLPCLCFPMGIMDASEVGLGLEYLKDNTNLFVGGLLLAIAVEHNLHKLRIALRVLLLTGVR 120
QY 121 PALLLGLFMLVTAFLSMWISNTATTAMVPIGHAVLEQLQSKKDVGGNNNPTFELOEE 180
DB 121 PALLLGLFVMTVAFSLMWSNTASTAMVPIAHAVLQELNNTQSNVEGSDNPTFELOEP 180
QY 181 CPQKEVTKL---DNG--QPUSA-PSEPTQKTQHHRFSQGLSICICYSASIGGIATLTG 234
DB 181 SPQKETSVDKDKNGQAPLPAPVLESGEHMTQRLRFSGMSLCVCYSASIGGIATLTG 240
QY 235 TPNVLVLOGVNSLFPQNGVNVNFASFAGFAPPMIILLALLWLOVLFLGVNFRKNFG 294
DB 241 TPNVLVLOGQWTSLFPQNPVNVNFASFAGFAPPMIILLALLWLOVLFLGVNFRKNFG 300
QY 295 FGEDEE--RKOAAFOVIKTYRLLGPMSPFAEKTVTVLVLLVLMFTREPGFPFGMDT 352
DB 301 IREQEHEQQRQAARVVIQTYRLLGPMSPFAEKAVALFVLVLLVLMFTREPGFFHGWNL 360
QY 353 VPANEKQSMASDGTVAIFSLVNFIIPSKIPGLMQDPKPKGKLKAPPAILTWKTNDKM 412
DB 361 VFDASGRVMVSDGSASILGVFLFVPSKIPGLTQDPDPNPGRLKAPPALLNKLNVKKM 420
QY 413 PWNIVLLGGGFALAKGSEGLSEWLGDKLTPLQHIPPSATAVILCLLAIETECTSNV 472
DB 421 PWNIVLLGGGFALAKGSEGLSQWGNKMLQHVPPPAVFIICLLVATTECTNSA 480
QY 473 ATTTFLPLILASMAQAICLHPLYVMLPCTLLAASLAFMLPVATPPNAIVFSFGGLKVSMA 532
DB 481 ATTTLLPLILASMAQAICLHPLYVMLPCTLLASSLAFMLPVATPPNAIVFSFGGLRVSDMA 540
QY 533 RAGFLNMIIGVLAITLSINWSPIFKLDTFPPSWAHNS-TSQCLLNPSTVP 584
DB 541 RAGIMLNIIGVLTMLAINSGVPMFQLHTFPPSWAHNSNTHCLASPPPTAPSP 593

RESULT 2
JC7911
Na+/sulfate cotransporter NaCT - human
C:Species: Homo sapiens (man)
C:Date: 31-Mar-2003 #sequence_revision 31-Mar-2003 #text_change 14-Jul-2003
C:Accession: JC7911
R:Inoue, K.; Zhuang, L.; Ganapathy, V.
Biochem. Biophys. Res. Commun. 299, 465-471, 2002

A:Title: Human Na+-coupled citrate transporter: Primary structure, genomic organization,
A:Reference number: JC7911; MUID:22334959; PMID:12445824
A:Accession: JC7911
A:Molecule type: mRNA
A:Residues: 1-568 <INO>
A:Cross-references: GB:A151833
C:Comment: This transporter classified as a tricarboxylate transporter represents the first
cellular entry of citrate by a process energized by the electrochemical Na+ gradient
or the synthesis of fatty acid and chol 17p12-13
C:Genetics:
A:Gene: nact

Query Match 53.0%; Score 1622; DB 2; Length 568;
Best Local Similarity 53.2%; Pred. No. 2.2e-123;
Matches 309; Conservative 108; Mismatches 132; Indels 32; Gaps 6;
QY 1 MATCWPAALWAYRFLYIVLCPLPILPLIVOTKEAYCAYSIIMALLWCTEALPLAVTA 60
DB 1 MASALSYVSKFKSFVILFVTPLLLPLVILMPAKFVRCAYVILMAIYWCTEVPLAVTS 60
QY 61 LPPIVLPLMGIMDASEVCIEYFKDTNLLFVGLMVAIAVEHNLHKLRIALQVLLIIGVR 120
DB 61 LMPVLLFPQLIDSRQVCVQYMKDNTNMLFLGGLIVAVAVERNLHKLRIALTLLMVGAK 120
QY 121 PALLLGLFMLVTAFLSMWISNTATTAMVPIGHAVLEQLQSKKDVGGNNNPTFELOEE 180
DB 121 PARLMLGFGVTAFLSMWISNTATTAMVPIVEAILQQMEATSAATEAG-----LELVDK 175
QY 181 CPQKEVTKLDNGQPVSAPESEPTQKTQEHHRFSQGLSICICYSASIGGIATLTGTPNLV 240
DB 176 GRAKE---LPGSQVIFEGPILGQOBDQERKRLCKAMTLCICVAASGTGTTATLTGTGNV 232
QY 241 LOGOVNSLFPQNGVNVNFASFAGFAPPMIILLALLWLOVLFLGVNFRKNFGFGEDEE 300
DB 233 LIGQWNELEPDSKDLNVNFASFAGFAPPMIILLALLWLOVLFLGVNFRKNFGSGGLSK 292
QY 301 ERKQAAFOVIKTYRLLGPMSPFAEKTVTVLVLLVLMFTREPGFPFGW-----GDTV 353
DB 293 KNEKAALKVLOEYRKLGPLSPAEINVLICPFLVILFSDPDGPMGMLTVAWVEGETK 352
QY 354 FANEKQSMASDGTVAIFSLVNFIIPSKIPGL-----MQDPKPKGKLKAPPAILTWKT 407
DB 353 Y-----VSDATVAIFVATLFLFVPSQPKFNPFRSQTEEBEETP---FYPPPLLDWKV 401
QY 408 VNDKMPWNIIVLLGGGFALAKGSEGLSEWLGDKLTPLQHIPPSATAVILCLLAIET 467
DB 402 TOEKVPWGVLLGGGFALAKGSEGLSVWKGMBELHVPVPAITLILSLVAVFTE 461
QY 468 CTSNVATTTFLPLILASMAQAICLHPLYVMLPCTLLAASLAFMLPVATPPNAIVFSFGGLK 527
DB 462 CTSNVATTTFLPLILASMAQAICLHPLYVMLPCTLLSASFAFMLPVATPPNAIVFTYGLK 521
QY 528 VSDMARAGFLNMIIGVLAITLSINWSPIFKLDTFPPSWAH 568
DB 522 VADMVKTGVIMNIIGVFCVFLAVNTWGRAIFDLDFPDWAN 562

RESULT 3
A47714
Na+/sulfate cotransporter, renal - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 02-Jun-1994 #sequence_revision 02-Jun-1994 #text_change 09-Jul-2004
C:Accession: A47714
R:Markovich, D.; Forgo, J.; Stange, G.; Biber, J.; Murer, H.
Proc. Natl. Acad. Sci. U.S.A. 90, 8073-8077, 1993
A:Title: Expression cloning of rat renal Na+/SO4(2-) cotransport.
A:Reference number: A47714; MUID:93376745; PMID:7690140
A:Accession: A47714
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-595 <MAB>
A:Cross-references: UNIPROT:Q07782; GB:L19102; NID:g310182; PID:AAA41677.1; PID:g310183
C:Superfamily: sodium/sulfate cotransporter

Query Match 41.2%; Score 1259; DB 2; Length 595;
Best Local Similarity 42.3%; Pred. No. 5.7e-94;
Matches 254; Conservative 123; Mismatches 176; Indels 48; Gaps 9;

QY 9 WAYRFLVILCPILFLLPLVOTKEAYCAYSIILMALLWCTEALPLAVTALFIVLFP 68
DB 8 FYRFLVLLVFTVLLPLPLIRKEAECAVILFVIATFWITEALPLSITALLPLMFP 67
QY 69 LMGIMDASEVCIEYFKDNTNLFVGLMVAIAVEHNLHKRIALQVLLIIGVRPALLLGF 128
DB 68 MEGIMSSHTVASAYFKDFHLLIGVICLSTSEKNLHKRIALRMVMGVNPAWLIGF 127
QY 129 MLVTAFLSMWISNTATTAMVPPIGHAVLEQLQSGSKOIVE-----GGNNNPTF 175
DB 128 MSSTAFLSMWISNTSTAAMVPIEVAQAQITSABAEAEATQMTYFNESAAGLEVDETI 187
QY 176 ELQECPCQKEVTK-----LNGQVPASBPRTQ-----KTEHHRFSQGLSLICY 222
DB 188 IQQETNRERKTKPALGSSNDKKG-VSSKMETEKNTVTGAKYRSKDHMMCKMLCIAY 246
QY 223 SASIGGIATLTGTPNLVLOGVNSLFPONGVNVNFASFVGFAPFTMIILLLLAWLQV 282
DB 247 SSTIGLTIITGTSNLFSEHFNTRY-PDCRLNFGSWFLSPFVAVILLLLSWINLOW 305
QY 283 LFLGVNFRKNGFGGEEBERKQAAFOVIKTOYRLGLGPMSPFAEKTIVTLFVLLVLMFTRE 342
DB 306 LFLGVNFRKNGFGGEEBERKQAAFOVIKTOYRLGLGPMSPFAEKTIVTLFVLLVLMFTRE 342
QY 343 PGFFPGWGTVPANKEGQSMASDGTVAIFISLVMFIIPSKIPGLMQDPKKPKLAP--- 399
DB 366 PGFVTGM-SVLFSYEPG--YVTDSTVALVAGILFLFLPA-----KLTQMTSTGDI 413
QY 400 -----PAILTWTNDKMPNIVILGGSEFALAKSGSEGLSEWLGDKLTPLQHIPPSAT 454
DB 414 IAFDYSPLITKBFOSFMPWDIILVGGFALADCCQVSGSSWIGSKLSPGSLFVWLI 473
QY 455 AVILCLIAIETECTNSVATTLFPLILASMAQAICLHPLVYMLPCTLAASLAFMLPVAT 514
DB 474 ILISSLIVTSLTEVASNPATITILFPLSPLAEAHVNPHLHLLSTICTSFAPLLPVAN 533
QY 515 PPAIVFSGGLKVDMARAGFLNIIIGVLAITLSINSWSPIPKLDTFPPSWAHNTSQC 574
DB 534 PPAIVFSGYGLKVIDMVKAGLVNGLVAVVLMGMFTWIBPMFNLHEYPSPADIVNQT 593
QY 575 L 575
DB 594 M 594

RESULT 4
S43561
YCR37C homolog K08E5.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 07-Sep-1994 #sequence_revision 10-Nov-1995 #text_change 24-Nov-1999
C:Accession: S43561
R:Kershaw, J.
Submitted to the EMBL Data Library, March 1994
A:Reference number: S43561
A:Accession: S43561
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-526 <KER>
A:Cross-references: EMBL:Z30974
C:Genetics:
C:Superfamily: sodium/sulfate cotransporter

Query Match 33.9%; Score 1037.5; DB 2; Length 526;
Best Local Similarity 40.9%; Pred. No. 4.2e-76;
Matches 217; Conservative 108; Mismatches 156; Indels 49; Gaps 8;

QY 38 CAYSIIIMALLWCTEALPLAVTALFPIVLPPLMGIMDASEVCIEYFKDNTNLFVGLMVA 97
DB 45 CAYCVCIIVYMSEVMPPLAVTALFPIVLPPLVGVLDANTTAKSYMNDTNLFVGLMVA 104
QY 98 IAVEHNLHKRIALQVLLIIGVRPALLLGFMLVTAFLSMWISNTATTAMVPPIGHAVLE 157
DB 105 AAVEKCDLHERVALSVLRVCGSBPKIMLGFWTMTALLSSGFSISNTATTAMVPPIQSVVQ 164
QY 158 QLQSGKDKVEGGNNNPTFELQECPCQKEVTKLDNGQVPASBPRTQKTQEHHRFSQGLS 217
DB 165 QL-----ISSFQHHPT-----NGERGLGCK-----KMATGLV 192

DB 37 CAYCVCIIVYMSEVMPPLAVTALFPIVLPPLVGVLDANTTAKSYMNDTNLFVGLMVA 96
QY 98 IAVEHNLHKRIALQVLLIIGVRPALLLGFMLVTAFLSMWISNTATTAMVPPIGHAVLE 157
DB 97 AAVEKCDLHERVALSVLRVCGSBPKIMLGFWTMTALLSSGFSISNTATTAMVPPIQSVVQ 156
QY 158 QLQSGKDKVEGGNNNPTFELQECPCQKEVTKLDNGQVPASBPRTQKTQEHHRFSQGLS 217
DB 157 QL-----ISSFQHHPT-----NGERGLGCK-----KMATGLV 184
QY 218 LCICYSASIGGIATLTGTPNLVLOGVNSLFPONGVNVNFASFVGFAPFTMIILLLLAW 277
DB 185 LSIICFAANIGTGATGTPSNLMLGQSLALFPKVDGSLNTVTWTFAYPLMLLCLFVAV 244
QY 278 LMLQVFLGVNFRKNGFGGEEBERKQAAFOVIKTOYRLGLGPMSPFAEKTIVTLFVLLVVL 337
DB 245 MTLVSFFL-----RDAPKDEAVTEMLKTRNELPRMTYAEKSVFVFCILLSL 293
QY 338 WFTREPFGFPFGWGTVPANKEGQSMASDGTVAIFISLVMFIIPSKIPGLMQDPKKPKL 397
DB 294 WVRNPGVWPGFG-VFF-KKG--AYTDATSAMIVAFLLFVLPSERPDLATVIKKE-DLK 347
QY 398 APPAILTWTNDKMPNIVILGGSEFALAKSGSEGLSEWLGDKLTPLQHIPPSATVI 457
DB 348 KRGCLMDMKTMQETFPMSVLLGGFALAGVKEGSLGSLGSLNLSIEHLPLWLQL 407
QY 458 LCLIAIETECTNSVATTLFPLILASMAQAICLHPLVYMLPCTLAASLAFMLPVATPPN 517
DB 408 TMLTAMVITNICTSVTASIFVPIVATLAQRAGHPPFTMLPTTLASSFAIFVGTGTPN 467
QY 518 AIVFSGGLKVDMARAGFLNIIIGVLAITLSINSWSPIPKLDTFPPSWA 567
DB 468 AIVFSGMWKVSMDAFVGGIISLELLVLTVLYNNSIAYLTLPLEFPTWA 517

RESULT 5
G88575
protein K08E5.2 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-May-2001
C:Accession: G88575
R:anonymous, The C. elegans Sequencing Consortium.
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
Science 282, 2012-2018, 1998
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/projects/C_elegans
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: G88575
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-534 <STO>
A:Cross-references: GB:chr_III; PIDN:CAA83225.1; PID:G3878357; GSPDB:GN00021; CESP:K08E5.
C:Genetics:
A:Gene: K08E5.2
A:Map position: 3
C:Superfamily: sodium/sulfate cotransporter

Query Match 33.9%; Score 1037.5; DB 2; Length 534;
Best Local Similarity 40.9%; Pred. No. 4.3e-76;
Matches 217; Conservative 108; Mismatches 156; Indels 49; Gaps 8;

QY 38 CAYSIIIMALLWCTEALPLAVTALFPIVLPPLMGIMDASEVCIEYFKDNTNLFVGLMVA 97
DB 45 CAYCVCIIVYMSEVMPPLAVTALFPIVLPPLVGVLDANTTAKSYMNDTNLFVGLMVA 104
QY 98 IAVEHNLHKRIALQVLLIIGVRPALLLGFMLVTAFLSMWISNTATTAMVPPIGHAVLE 157
DB 105 AAVEKCDLHERVALSVLRVCGSBPKIMLGFWTMTALLSSGFSISNTATTAMVPPIQSVVQ 164
QY 158 QLQSGKDKVEGGNNNPTFELQECPCQKEVTKLDNGQVPASBPRTQKTQEHHRFSQGLS 217
DB 165 QL-----ISSFQHHPT-----NGERGLGCK-----KMATGLV 192

Db 400 LACSLNSTSKMPFFVMOIILSIVVVMTFSTNSATASIFIPISFKMAEAVGAHPLYF 459
Qy 497 MLPCTLAASLAFMLPVPAPNAIVFSGGLKYSVDNARAGFLNIIIGVLAITLSINSWSIP 556
Db 460 SIPTAIGPSFSPMLPWATANAIVETKIRIMIDWVSCGVFNIPICIAITAINMTWAPW 519
Qy 557 IPKLOTFFPSWA--HS-----NTSQC 574
Db 520 LFNMGTYPDYALRHATNMTGNSQC 544

RESULT 8

Cy9980
Hypothetical protein SA1732 [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: C89980
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogura, A.; Mizutani-U, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.; Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: C89980
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-520 <KUR>
A:Cross-references: UNIPROT:Q99SX1; GB:BA000018; PID:g13701709; PIDN:BA843002.1; GSPDB:C89980
A:Experimental source: strain N315
C:Genetics:
A:Gene: SA1732
C:Superfamily: sodium/sulfate cotransporter

Query Match 29.7%; Score 907.5; DB 2; Length 520;
Best Local Similarity 37.0%; Pred. No. 1.4e-65;
Matches 198; Conservative 97; Mismatches 159; Indels 81; Gaps 10;
Qy 22 IPLLPL----PLIVQTKBAYCAYSIIIMALLWCTEALPLAVTALPPIVLPFLMGMIMDASE 77
Db 42 LPLLTLFPHPQDLPKGVVLAITLWATWITAIPIATSLPIVLPPLGLHILTEQ 101
Qy 78 VCIEYFKDTNLIIFVGLGMVAIAVEHNLHKLIALQVLIIGVRPALLLGLFMLVTAFLSM 137
Db 102 VSSEYENDIIFLFGGFTLAIAERWNLHTRVALTIINLIGASTSKILLGFMVATGFLSM 161
Qy 138 WISNTATTAMVPIGHAVLEQLOGSKDVEGNNPTPELOECQKVEKTLKDNQPVSA 197
Db 162 FVSNAAVMIMIPIGLAIKEAH-----DLOEANTNQ----- 194
Qy 198 PSEPTQKTQEHHRFSQGLSLCICYSASTGGIATLTGTPNLVLQGVNSLPFGNGNVN 257
Db 195 -----STQKEKSLVLAIGVAGTIGGLTGLTTPPLIILKGQWQH-----GHIS 241
Qy 258 FASWFGFAPPTMIILLLAWLQVLFLGVNFRNFGFGEGERKQAAFOVIKTYRL 317
Db 242 FAKMIVGPTIVILGIGITWLYR-----VAFRHLKYLPGQ-----TLIKQLDEL 290
Qy 318 GPMFAEKTIVTVLVLLVVLVFTREPGFPGMGDTVPANEGQ--SMASDGTVAIFSLV 375
Db 291 GKMKEEVKQVQTFVLASLLWITRE-----FLKKWEVTSVADGTIAIFISIL 339
Qy 376 MFIIPSKIPGLMQDKPKGLKAPPAITLTKVNDKMPNIVILGGGFALAKGSEQSL 435
Db 340 LFIIPAK-----NTEKHRIIDWE-VAKELPWGVLLIPGGGLALAKGISESL 386
Qy 436 SEWLGDKLTPLOHIPPSTAVILCLLIAIFTECTSNVATTTLFPLILASMAQAIICLHPY 495
Db 367 AKWLGEQLKSLNGVSPILVIVITFVLFLEVTSTNTATMTILPILATLSVAVGVHFL 446
Qy 496 VMLPCTLAASLAFMLPVPATPPNAIVFSGGLKYSVDNARAGFLNIIIGVLATLSI 550
Db 447 LMAPAAMAANCAYMLPVGTTPNAITFGSGKISIKQMASVGFWNLSIAIIILVV 501

RESULT 9

A88546
protein R107.1 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: A88546
R:anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/c_elegans/ and www.ganger.ac.uk/Projects/C_elegans/
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: A88546
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-539 <STO>
A:Cross-references: UNIPROT:P32739; GB:chr_III; PIDN:CAA78468.1; PID:g3879033; GSPDB:GN01
A:Note: predicted using GeneFinder
C:Genetics:
A:Gene: R107.1
A:Map position: 3
C:Superfamily: sodium/sulfate cotransporter

Query Match 27.2%; Score 830.5; DB 2; Length 539;
Best Local Similarity 32.7%; Pred. No. 2.5e-59;
Matches 190; Conservative 126; Mismatches 198; Indels 67; Gaps 12;
Qy 15 LIVLCLPIFLPLPIVQTKBAYCAYSIIIMALLWCTEALPLAVTALPPIVLPFLMGMIMD 74
Db 13 LVLVLLGP--LVAVPLLPFGPEYRCLFSIIFLSTYWIGEAFFPIGVTSLPPLALYPLIQIVP 70
Qy 75 ASEVCIEYFKDTNLIIFVGLGMVAIAVEHNLHKLIALQVLIIGVRPALLLGLFMLVTAFL 134
Db 71 SKQISPVFKDSIVFLMCTLIMAMAVEATGLHRIALKLLTKVGAQP----- 118
Qy 135 LSMWISNTATTAMVPIGHAVLEQLOGSKDVEGNNPTPELOECQKVEKTLKDNQPV 194
Db 119 VSPFVSDTACTALMCPATAVALLMSMSDAVQHLKEDHRKPK--PPDDATVAEKLRIIDWTP 177
Qy 195 VSASEPRTQKTQEHHRFSQGLSLCICYSASIGGIATLTGTPNLVLQGVNSLPFGNGNV 254
Db 178 QDA-----GFCALILACAHASLIGGTAITITSTGPNLVFRENTHKRYPEQOV 224
Qy 255 VVNFASWFGFAPPTMIILLLAWLQVLFLGVN-ERKNVGFGEGER--KQAAFOVIK 312
Db 225 TMTYLQWVFAIPPMFVYLLASYIILVCYFMGPSTFARWFERPERSKEEHLKLEKNIQ 284
Qy 313 QYRLIGPMSFAEKTIVTVLVLLVVLVFTREPGFPGMGDTVPANEGQSMASDGTVAIFI 372
Db 285 MYEDLDGVNKGESVFFVILLIGSWISRODPGFTPGWGDLL----PHRNFI SDSVSGVLI 340
Qy 373 SLVNFIIIPSKIPGLMQDKPKGLKAPPAITLTKVNDKMPNIVILGGGFALAKGSEQ 432
Db 341 SCILFVWP-----KDPFDPIDPMAP--ILKWTDMKSKFSWCTLLIGAGYAISEGVDK 391
Qy 433 SGLSE-----WLGDKLTPLOHIPPSTAVILCLLIAIFTECTSNVATTTLFPLILA 483
Db 392 SGLSRLISCGMKNIFVGMSSLPLQ-----LVTVTIIVIMTEFASNVSTGSIPIPSL 443
Qy 484 SMAQAIICLHPYVLMPLPCTLAASLAFMLPVPATPPNAIVFSGGLKYSVDNARAGFLNIIIGV 543
Db 444 GVAESMGVHLYALPPTVACSFAPMLPISTPPNAVVDYTKVSWEMIVCGFLINLTIACI 503
Qy 544 LAITLSINSWSIPFKLDTFPSWAHSNTSQCLNPSNSTVP 584
Db 504 LITSLNMTWTYFISLNIFF-----ENIVISSENSVYP 537

RESULT 10

S30871
hypothetical protein R107.1 - Caenorhabditis elegans

A;Reference number: A71800; MUID:99120557; PMID:9923682
A;Accession: E71961
A;Status: Preliminary
A:Molecule type: DNA
A;Residues: 1-548 <ARN>
A;Cross-references: UNIPROT:Q9ZML8; GB:AE001458; GB:AE001439; NID:g4154713; PIDN:AAD05784
A;Experimental source: strain J99
C;Genetics:
A;Gene: jhp0200
C;Superfamily: sodium/sulfate cotransporter

Query Match 25.7%; Score 785; DB 2; Length 548;
Best Local Similarity 34.7%; Pred.No.1.2e-55;
Matches 181; Conservative 100; Mismatches 172; Indels 68; Gaps 14;

Qy 43 ILMALLWCTEALPLAVTALFPVILFPLMGINDASEVCIEYFKDTNLIFVGGLMVATAVEH 102
Db :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
Qy 68 VLGMIGMMTEAIDLPAATALLPLVLNFVSVDQFSVSASYASPIIPLFWGGFFILSMQK 127
Db :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
Qy 103 WNLHKRIALQVLLIIIVRPALLLGLFMVLTPFLSNWISNTATTAMVPIGHAVLEQLQGS 162
Db |||||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
Qy 128 WNLTRIALSIILLVGSPPRLILGLFMIAITGFLSNWWSNTATAVMMFPVGMVSLQLV--- 184
Db |||||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
Qy 163 KKDVEGGNNPTFELOECPOKEVTK-----LDNGQPVSAPSEPTKTQEHHRFS 213
Db :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
Qy 185 -AKLVGKNASNAPYQKE----ETKAHGGINMSNVHKGDIAQVIQEK--TIYRTNFS 237
Db :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
Qy 214 QGLSICYSASIGGIATLTGTTLNVLIOGVNSLPQNQNV-VNPASFPGFAFTMIIL 272
Db |||||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
Qy 238 ICLMLGIAYSASIGSLGTLIQTTPNALLAGYMKTAF---NTEIDPAQWMVFCTPLAFIM 293
Db |||||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
Qy 273 LLLAWLMQLVLFLGNFRKFNFGCEGEERKQAQFVLTKQYRLLGPMSPAEKTVTVLFV 332
Db :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
Qy 294 LLLAWLLTVIFFELKIKEIQ-----VKSEDKLRRLSQAEISVGVIPI 342
Db :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
Qy 333 LLVVLFNTRPFGPFPGWG---DITVFANEKGOSMASDGTVAIFISLVMFIIPSKIPGLMQD 389
Db :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
Qy 343 LASLGWIFLDT-ILKSWGIKIDKI-----DSVIAMGVSAALLFILPANHQG--- 386
Db :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
Qy 390 PKPCKGLKAPAAILTWKTYNDKMPWNIVILLGGGFALAKGSEQSLSEHLGDKLTPLOHI 449
Db :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
Qy 387 -----DRLLDW-GVAKKLWDVLLFLFPGGLLSAQFSKTLGSLMIGHLVSGFSHL 435
Db :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
Qy 450 PPSATAVILCILIAIFTESNVAITTLFILPASMAQAI---CLHPLVYMLPCTLAASIL 506
Db :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
Qy 436 PILFIIWVTLWLVIPLFELITSNTATAAAFLPVIGGVANGMGYESHQSLLLTIPVALSATC 495
Db :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
Qy 507 AFMLPVATPNNAIVSFPGGLKVSDMARAGFLNIIGVLAIT 547
Db :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
Qy 496 AFMLPVATPNNAIAYGSGYKITDMIKAGLWLNLVGVVLIS 536
Db :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:

RESULT 12
F64546
sodium-dependent transporter - Helicobacter pylori (strain 26695)
C;Species: Helicobacter pylori
C;Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 09-Jul-2004
C;Accession: F64546
R;Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenney, son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L. Nature 388, 539-547, 1997
A;Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.N.
A;Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A;Reference number: A64520; MUID:97394467; PMID:9252185
A;Accession: F64546
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A;Residues: 1-552 <TOM>
A;Cross-references: UNIPROT:O25003; GB:AE000541; GB:AE000511; NID:g2313299; PIDN:AAD07282
C;Superfamily: sodium/sulfate cotransporter

Query Match 25.5%; Score 781; DB 2; Length 552;

Best Local Similarity 34.5%; Pred. No.2.6e-55;
Matches 182; Conservative 100; Mismatches 165; Indels 80; Gaps 16;

Qy 43 ILMALLWCTEALPLAVTALFPVFLPMLGMDASEVCIEYFKDNLILFVGSLMAVAIAVEH 102
Db :|::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||:
72 VLMIGWMTTEAIDLPATALLPLVLFSVFSDVPQASVSSTASPIIFLPMGFGLSNMQK 131
Qy 103 WNLHKRIALQLVLLIIGVRPALLLGLFMLVTAFLSMWISNTATTAMMPVIGHAVLE--QL 159
Db :|::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||:
132 WNLHTRIALSIIILLVGTSPRLILGFMNATGFLSMWTNTATAVMMLPVGMVSLQLVAKL 191
Qy 160 QGSKKDVGGNNNTFFELQECPKEVKTLONG--QPVSAPSEPRKTQTQB-----HHRF 212
Db :|::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||:
192 VG-KEDASNWHQ-----KBEITKAHGSGIMSNIVHKGKDITQVOEKTTTYRTNF 240
Qy 213 SOGLSCLICYSAGGIATLTGTPNLVLQGVSNSLPQNGNV-VNFASNFGFAFTMI 271
Db :|::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||:
241 SICMLGIAYNASIGSLTGTLTPNALLAGMYKTAF----NIETDAQMWWFTGPLAFI 296
Qy 272 LLLLAWLMLQLVLFUGVNFRKNFGEGEEERKQAFOVIKTQVRLLEGMSFAEKTVTULVF 331
Db :|::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||:
297 MLILAULLLVTFPLKIETPG---GKE-----VIRVELKKLGRLSQAEISVGIIIF 345
Qy 332 VLLVVUHFTRPPGFPG-----WG-----DTVPANEKGQSMDSGTVAFISLNWFIPDSKI 383
Db :|::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||:
346 ILASLGV-----IFLGVMLKSWGVIDKI-----DSVIAMGVSAALLFILPANH 388
Qy 384 PGLMQDPKKPKGLKAPPAILTWTKNDKPWNIVILLGGGFALAKGSEQSLSEWLGDKL 443
Db :|::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||:
389 QG-----DRLDIW-GVAKKLPMVDVLLLFGGGLAUSAQSKTGLSLWIHLV 433
Qy 444 TPLQHIPPSSATAVILLCILAIIFTESNSTNAVTTTLFLPILASMAQAICL---HPLYVMMLPC 500
Db :|::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||:
434 SGFSLHPILFIWMVTLWVILEITTSNTATAAAFLPVIGGVAMGMYENHQSLLLTIPTV 493
Qy 501 TLASLAFMLPVATPPNAIVPSFGCLKSVDMARAGFULNIIGVLAIT 547
Db :|::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||:
494 ALSATCAFMLPVVTPPNNAIYGSGYVKITDIKAKGLNLNLVGVVLIS 540

RESULT 13
I64080
probable membrane protein HI0608 - Haemophilus influenzae (strain Rd KW20)
C:Species: Haemophilus influenzae
C>Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 09-Jul-2004
C:Accession: J64080
R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.
J.; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman,
J., D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Flohrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A.:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A:Reference number: A64000; MUID:95350630; PMID:7542800
A:Accession: J64080
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-461 <TIGR>
A:Cross-references: UNIPROT.Q57486; GB:U22743; GB:L42023; NID:g1573597; PIDN:AAC22267.1;
C:Superfamily: probable transporter MJ0672

Query Match 18.2% Score 556; DB 2; Length 461;
Best Local Similarity 28.5%; Pred. No.3.4e-37;
Matches 158; Conservative 95; Mismatches 177; Indels 124; Gaps 16;

Qy 15 LIVLCLBIFILPLPLIVQTKEANCYSII-LMALWCETEALPLAVTALFPVFLPMLMGIM 73
Db :|::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||:
15 IFILDIVLFFVLLNLVPEPKANSGLALLAFIAVLWSLEAHVTTITALLPVLLVALGLV 74
Qy 74 DASEVCIEYFKDNLFPVGLMVAIAVEHWNLKHRIAQLVLLIIGVRPALLLGLFMLVTA 133
Db :|::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||:
75 STQALVGFADPTIFLFFGGFSATALHIQOKLKANKLMARGNFLTAVILFELITA 134
Qy 134 FLNSWISNTATTAMMPVIGHAVLEQLQSGSKDVEGGNNNPTELQECBPQEVTKLDNGO 193

```

Db      135 FLSSMTMENTATAAMMLPLAMGILLSQL-----DREKDHNTYVFL----- 173
Qy      194 PVSAPSPRTQTQOEHRRFSQGLSLCICYSASIGSIATLTGTTPNLVLQGQVNSLFPQNG 253
Db      174 -----LGTAYSASIGMGTLVGSPPPNAIVASNLN----- 202
Qy      254 NVNPFASFQGFAPPFTMIILLALLWLMLQVLV---LGVNFRKNFGFGEGEREERKOAQFOVI 310
Db      203 --LTFSDMLWVGLPIMIIILPLMIGILVIIIPKKLHLNFETFF-----ENIE----- 247
Qy      311 KTYRLLLGPSFAEKTTVVLFVLLVNLWFTRPECFPGWGDTVFANEKG--QSMAS-DGT 367
Db      248 -----MNPMRI-----LTFIIFPVIALTW-----IFSCKINPFISGLGLQKNIAFSOSI 292
Qy      368 VAFISLVMEFIIPSIPCLMQDPKPKGLKAPPAILTWKTVDNKMPNMIVILLGGGFALA 427
Db      293 VALLAAIVI-----CSTGVASKWKQTSNDWGMVLMPFGGLTLS 331
Qy      428 KGSEQSGLSWLDKLPPL---QHPPSATAVILCLLIAIF----TECTSNVATTTLFLP 480
Db      332 AVLKDSGASKILADSIIVFMIDGQHF-----YLIGLLVAEFIPLTEPTNSTASAALLVP 385
Qy      481 ILASMAQAICLHPLYVMLPCTLAASLAFMLPVATPPNAIVPSFGLKVSDMARAGFLNNI 540
Db      386 IFISIAQSGLMPEIGLALITIGASCAPMLPVATPPNAIVFGSQVKOSEMVKVGFLLNL 445
Qy      541 IGVLAI-TLSINSW 553
Db      446 VCVVVIATMGYMF 459

RESULT 14
F70302
transporter (Pho87 family) - Aquifex aeolicus
C:Species: Aquifex aeolicus
C>Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004
C:Accession: F70302
R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham
V.
Nature 392, 353-358, 1998
A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolic
A;Reference number: A70300; MUID:98196666; PMID:9537320
A;Accession: F70302
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-446 <AQF>
A;Cross-references: UNIPROT:O66449; GB:AEO00670; NID:g2982779; PIDN:AAC06409.1
A;Experimental source: strain VF5
C;Genetics:
A;Gene: trns
C;Superfamily: probable transporter MJ0672

Query Match          17.9%; Score 548; DB 2; Length 446;
Best Local Similarity 31.2%; Pred. No. 1,5e-36;
Matches 157; Conservative 74; Mismatches 163; Indels 110; Gaps 14

Qy      42  ILMALLWCTEALPLAVTALFPVILFPLMGMDASEVCIEYFKDTNI-LFVGGLMVAJAV 100
Db      34  LLLAAILWITEALPLPVTALLIPVSGVLLGVFDV-KTALSFYAHPLIFLPFGGFVALVAL 92
Qy      101 EHNHAKRIALQVLLIIGVR--PALLLGFWLVTAFLSMTSINTATTAMVMPIGHAVLEQ 158
Db      93  SKYQIDEYIAHKIVSAQGFPLSPVLL--MLATSLISMNISNTTATMLPLAIGLAG 150
Qy      159 LOGSKKDVEGGNNNPTELQECPQKEVTKLDNGQPVSAPSEPRTOXTQEHHRFSQGLSL 218
Db      151 V-----RETEREKVPFPVLL 165

Qy      219 CICYASISIGGIATLTGTTPNMLVLQGVNSLFPONGNVNPFASWFGFAPPTMIILLALLAWL 278
Db      166 GRAYSASVGIGITLVGSPPNGIAAGILGLSFP-----DWLKFGIPVPLIFPLPLF- 215

```

```
QY 279 WLQVLFLGVNFRKFGGEBERKQAAFOVIKTOYRLGPMSPAEKTTVTFLVLLVLM 338
Db 216 --AIIFL--VFPT-----SDLKVERVQEK-----FETPQRLVLLIFL----- 252
QY 339 FTREPGFFPGMDTVFANEKGQSMASDGTVAIFISLVNFIIPSKIPGLMQDPKPKGLKA 398
Db 253 FTALAWIFSKKIAPIFEVKK-----YFDTVALLAVVLFIF----- 289
QY 399 PPAIITWTKMKNWKNIVILLGGGFALAKGSEQSLSEWLGDKLTPLQHPPTSPATAVIL 458
Db 290 --RLDMDRDVKEGVSNGTLLIFGGGIALSGIMKKTGTAKFISQBLVDVLHLGLTFLLT 347
QY 459 CIIIAIF-TECTSNVATTTFLPILASMAQAICLHPLYVMLPCTTLAASLAPMLPVATPPN 517
Db 348 IVLFVIFLTELMSNTATTALAPILFSTAQWIGKPEMLVIPAAVAASCAFMPLPVATPPN 407
QY 518 AIVFSFGGLKVSMDMARAGFLNII 541
Db 408 AIVYGTGYIKSQMMRVGLILNIV 431

RESULT 15
B82510
transporter, NadC family VCA0025 [imported] - Vibrio cholerae (strain N16961 serogroup C
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: B82510
R:Heidelberger, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, B
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: B82510
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-462 <HEI>
A:Cross-references: UNIPROT:Q9KNE0; GB:AE004346; GB:AE003853; NID:g9657401; PIDN:AAF9593
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VCA0025
A:Map position: 2
C:Superfamily: probable transporter MJ0672
```

```
Query Match 17.2%; Score 525.5; DB 2; Length 462;
Best Local Similarity 29.7%; Pred. No. 1e-34;
Matches 153; Conservative 74; Mismatches 163; Indels 125; Gaps 13;

QY 42 IIMALLMCTEALPLAVTALPPIVLFPPLMGIMDASEVCIEYFKDTNILFVGGLMVAIAVE 101
Db 52 LAFTAVLWLTALHVTVTAILVPVMAVFFGIFETQAAALNNFANSIIFLFGGFALAAAMH 111
QY 102 HNNLHKRIALQVLLIIGVRPALLLLGFWLVTAFLSMISNTATTAMVPICHAVLEQLOG 161
Db 112 HOGLDKVIADKVLAMAQKMSVAVFMLFGVTALLSMISNTATAAMMLPLVLGVL----- 166
QY 162 SKDVEGGNNPTPELOECQKEVTKLDNGQPVSPAPSEPRTKQTOEHHRFSQGLSLCIC 221
Db 167 SKVDADKQRTSYTVFL-----LGVA 186
QY 222 YSASIGGIATLTGTPNLVQGVNSLFPQNGVNVNFASFVGFAPPTMIILLLLAMLWLQ 281
Db 187 YSASIGGIATLVGSPPPNAIAAEVG-----LSFTDMWKFGLPTAMMMLPMA----- 232
QY 282 VFLGVNFRKFGGEBERKQAAFOVIKTOYRLGPMSPAEKTTVTFLVLLVLMFTR 341
Db 233 -----IAILYFLKP-----TLNGMPEL-----DR 252
QY 342 EPGFFPGMDTVFANEKGQSMASDGTVAIF-ISLWMEIIPSKIPGLMQDPKPKGLKAPP 400
Db 253 AP-----VNW-----DKGVV-----TLGIFGLTVFLWIFSPINAAAGGKSFDTLVAG 298
QY 401 AILT-----WKTVNDKMPWNITVILLGGGFALAKGSEQSLSEWLGDKLTPL-QHIP 451
```

```
Db 299 AIIMLSFARVVHKEIQKTADWGVLLFGGLCLSNVLKQTGTSVFLANALSDMVSHMGI 358
QY 452 SATAVILCLLIIAIFTTECTSNVATTTFLPILASMAQAICLHPLYVMLPCTTLAASLAPMLP 511
Db 359 FVILVATFVVFELTEFASNTASAALLIPVFATVAEAFGMSPVLLSVLIAVAASCAFMPLP 418
QY 512 VATPPNAIVFSFGGLKVSMDMARAGFLNII--TGVL 544
Db 419 VATPPNAIVFASGHKQSEMRVGLYINIACTGLL 453
```

Search completed: June 30, 2005, 09:00:17
Job time : 19.5314 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2005, 08:28:39 ; Search time 78.9921 Seconds
(without alignments)
3805.323 Million cell updates/sec

Title: US-10-017-479A-4

Perfect score: 3058

Sequence: 1 MATCPALWAYRFYLIVLCL.....HSNTSQCLLNPSNVTVPGL 587

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3058	100.0	587	2	Q35055 rattus norv
2	2956	96.7	587	1	P70545 rattus norv
3	2837	92.8	586	1	Q13183 homo sapien
4	2392	78.2	592	1	Q13183 homo sapien
5	2340.5	76.5	593	1	Q28615 oryctolagus
6	2178	71.2	605	2	Q866R0
7	1971.5	64.5	622	2	O57661 xenopus lae
8	1967.5	64.3	624	2	Q642N5
9	1867	61.1	613	2	Q6P3L2
10	1862	60.9	613	2	Q803K7
11	1768.5	57.8	586	2	Q6AZR9
12	1624	53.1	568	2	Q86YT5
13	1610	52.6	568	2	Q6ZMG1
14	1536.5	50.2	572	2	Q67BT3
15	1513.5	49.5	572	2	Q8CJ44
16	1378.5	45.1	600	1	S133_MOUSE
17	1377.5	45.0	600	1	S133_RAT
18	1367.5	44.7	602	2	Q6EJG8
19	1364.5	44.6	602	2	Q6DJN7
20	1355	44.3	602	1	S133_HUMAN
21	1354	44.3	602	2	Q81VE1
22	1325	43.3	595	1	S131_HUMAN
23	1322	43.2	601	2	Q9W7I2
24	1316.5	43.1	583	2	Q6NY28
25	1316.5	43.1	583	2	Q6PE27
26	1311.5	42.9	596	2	Q6NMY4
27	1259	41.2	595	1	S131_RAT
28	1256	41.1	627	1	S134_HUMAN
29	1253.5	41.0	626	2	Q8N631
30	1253	41.0	601	2	Q7ZWL2
31	1248	40.8	625	2	Q8BZ82

ALIGNMENTS

32	1236	40.4	594	1	S131_MOUSE
33	1234	40.4	594	2	Q80YB5
34	1228	40.2	619	2	Q6L970
35	1184.5	38.7	520	2	Q8N8K4
36	1043	34.1	566	1	NAD3_CAEEL
37	1042.5	34.1	572	1	IND1_DROME
38	1042.5	34.1	590	2	Q7KUS6
39	1007	32.9	582	1	NAD1_CAEEL
40	968	31.7	896	2	Q7QIT2
41	937	30.6	552	2	Q8ENE3
42	930	30.4	450	2	Q75MH3
43	925	30.2	577	1	YK56_CAEEL
44	907.5	29.7	520	2	Q99SX1
45	907.5	29.7	520	2	Q7A4P8
46	907.5	29.7	520	2	Q6GFE0
47	906.5	29.6	520	2	Q8NV55
48	906.5	29.6	520	2	Q6G816
49	894	29.2	504	2	Q86B89
50	889.5	29.1	551	1	NAD2_CAEEL
51	877	28.7	302	2	Q68D44
52	873.5	28.6	505	2	Q811B6
53	871	28.5	562	1	IND2_DROME
54	866.5	28.3	546	2	Q65NG0
55	831.5	27.2	540	2	Q8LGR8
56	829.5	27.1	540	2	Q9MAW4
57	826	27.0	462	2	Q9F8H9
58	820.5	26.8	540	2	Q9F8H9
59	794	26.0	626	2	Q6AQ77
60	792	25.9	548	2	Q6ZPH7
61	785	25.7	548	2	Q9ZML8
62	781	25.5	522	2	Q250K3
63	775	25.3	524	2	Q6NKS9
64	773	25.3	510	2	Q8NT57
65	773	25.3	527	2	Q6M8D0
66	763.5	25.0	495	2	Q67T81
67	755	24.7	474	2	Q65MM0
68	754	24.7	587	2	Q8FU27
69	718.5	23.5	510	2	Q7UUK9
70	714.5	23.4	414	2	Q7PF67
71	659.5	21.6	440	2	Q81V67
72	644	21.1	450	2	Q8FM25
73	613	20.0	334	2	Q6ZNH7
74	565	18.5	464	2	Q8VTX1
75	565	18.5	464	2	Q9AED9
76	565	18.5	464	2	Q72QJ5
77	561	18.3	464	2	Q8F5L4
78	559	18.3	464	2	Q84F3
79	557.5	18.2	470	2	Q6SR88
80	556	18.2	461	1	Y608_HABIN
81	548	17.9	446	2	Q6C449
82	545.5	17.8	463	2	Q9CLQ0
83	537	17.6	451	2	Q7M8Y1
84	535	17.5	464	2	Q7ZGJ3
85	525.5	17.2	462	2	Q9KNE0
86	519.5	17.0	456	2	Q87Q94
87	518	16.9	463	2	Q65SS5
88	513.5	16.8	456	2	Q6LQW4
89	506.5	16.6	456	2	Q7MLG6
90	506.5	16.6	456	2	Q8D908
91	500.5	16.4	471	2	Q9K032
92	498	16.3	459	2	Q666I5
93	496.5	16.2	456	2	Q8ZHX0
94	495	16.2	459	2	Q8CZT3
95	494.5	16.2	471	2	Q9JVA3
96	492	16.1	461	2	Q7N822
97	490	16.0	474	2	Q8GF70
98	478.5	15.6	443	2	Q26881
99	468.5	15.3	487	2	Q9KSE3
100	464.5	15.2	471	2	Q6LSU9

```
RESULT 1
C35055 PRELIMINARY; PRT; 587 AA.
AC 01-JAN-1998 (TREMELrel. 05, Created)
DT 01-AUG-1998 (TREMELrel. 07, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Sodium/dicarboxylate cotransporter 1 (NA(+)/dicarboxylate
DE cotransporter 1) (Kidney dicarboxylate transporter) (SDCT1) (Organic
DE anion transporter 1) (OAT1).
GN Name=Nadc-1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=KIDNEY;
RX MEDLINE=98355911; PubMed=9691021;
RA Sekine T., Cha S.H., Hosoyamada M., Kanai Y., Watanabe N., Furuta Y.,
RA Fukuda K., Igarashi T., Endou H.;
RT "Cloning, functional characterization, and localization of a rat renal
RT Na+-dicarboxylate transporter.";
RL Am. J. Physiol. 275:F298-F305(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=KIDNEY;
RX MEDLINE=97373539; PubMed=9228014; DOI=10.1074/jbc.272.30.18526;
RA Sekine T., Watanabe N., Hosoyamada M., Kanai Y., Endou H.;
RT "Expression cloning and characterization of a novel multispecific
RT organic anion transporter.";
RL J. Biol. Chem. 272:18526-18529(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY;
RX MEDLINE=98361968; PubMed=9694847; DOI=10.1074/jbc.273.33.20972;
RA Chen X.Z., Shayakul C., Berger U.V., Tian W., Hediger M.A.;
RT "Characterization of a rat Na+-dicarboxylate cotransporter.";
RL J. Biol. Chem. 273:20972-20981(1998).
CC -1- FUNCTION: MAY MEDIATE ELECTROGENIC, SODIUM-DEPENDENT TRANSPORT OF
CC MOST KREBS CYCLE INTERMEDIATES, INCLUDING CITRATE, SUCCINATE,
CC ALPHA-KETOGLUTARATE, AND OXALOACETATE. TRANSPORT OF CITRATE WAS PH
CC SENSITIVE. IT MAY HAVE A SODIUM-SUBSTRATE STOICHIOMETRY OF 3:1. AT
CC LEAST ONE NA(+) SEEMS TO BIND TO THE TRANSPORTER BEFORE THE
CC SUBSTRATE.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN S3 SEGMENTS OF KIDNEY PROXIMAL
CC TUBULES AND IN ENTEROCYTES LINING THE INTESTINAL VILLI. ALSO
CC DETECTED IN LUNG BRONCHIOLI, THE EPIDIDYMIS, AND LIVER.
DR EMBL; AB001321; BAA28609.1; -.
DR EMBL; AF058714; AAC31165.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0015293; F:symporter activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006814; P:sodium ion transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR001898; Na/sul_symport.
DR Pfam; PF00939; Na sulph sym; 1.
DR TIGRFAMs; TIGR00785; dass; 1.
DR PROSITE; PS01271; NA SULFATE; 1.
KW Sodium transport; Symport; Transmembrane; Transport.
FT TRANSMEM 13 33 POTENTIAL.
FT TRANSMEM 53 73 POTENTIAL.
FT TRANSMEM 80 100 POTENTIAL.
FT TRANSMEM 114 134 POTENTIAL.
FT DOMAIN 123 126 POLY-LEU.
FT TRANSMEM 215 235 POTENTIAL.
FT TRANSMEM 264 284 POTENTIAL.
FT DOMAIN 272 275 POTENTIAL.
FT TRANSMEM 319 339 POTENTIAL.
FT TRANSMEM 367 387 POTENTIAL.
FT TRANSMEM 445 465 POTENTIAL.

FT TRANSMEM 477 497 POTENTIAL.
FT TRANSMEM 506 526 POTENTIAL.
FT TRANSMEM 535 555 POTENTIAL.
FT CARBOHYD 570 580 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 580 580 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 587 AA; 64142 MW; 844E47C77F89CDA CRC64;

Query Match 100.0%; Score 3058; DB 2; Length 587;
Best Local Similarity 100.0%; Pred. No. 3.4e-223; Indels 0; Gaps 0;
Matches 587; Conservative 0; Mismatches 0;

QY 1 MATCWPAALWAYFYLVLCPIFLPIFLPLIVOTKEAYCAYSIIIMALLMCTEALPLAVTA 60
DB 1 MATCWPAALWAYFYLVLCPIFLPIFLPLIVOTKEAYCAYSIIIMALLMCTEALPLAVTA 60
QY 61 LPPIVLFPPLMGINDASEVCIEYFKDNIILFVGLVMVAIAVEHNLHKLALQVLLIIGVR 120
DB 61 LPPIVLFPPLMGINDASEVCIEYFKDNIILFVGLVMVAIAVEHNLHKLALQVLLIIGVR 120
QY 121 PALLLLGFMVLTAFLSMWISNTATTAMVPIGHAIVLEQLQSGKDVGGNNNPTFELOE 180
DB 121 PALLLLGFMVLTAFLSMWISNTATTAMVPIGHAIVLEQLQSGKDVGGNNNPTFELOE 180
QY 181 CPQKEVTKLDNGQPVSAPEPRTKTQEHRRFSQGLSLCICYSASIGGIATLTGTTPLNV 240
DB 181 CPQKEVTKLDNGQPVSAPEPRTKTQEHRRFSQGLSLCICYSASIGGIATLTGTTPLNV 240
QY 241 LQGVNSLFPQNGVNVNPFASWFGFAPPTMIILLLAWLWQLVFLGNFRKNGFGEGEE 300
DB 241 LQGVNSLFPQNGVNVNPFASWFGFAPPTMIILLLAWLWQLVFLGNFRKNGFGEGEE 300
QY 301 ERKQAAFOVTKQYRLGPMGSAEKTIVLVLLVVLWFTREPPGPPGMDTVFANEKGQ 360
DB 301 ERKQAAFOVTKQYRLGPMGSAEKTIVLVLLVVLWFTREPPGPPGMDTVFANEKGQ 360
QY 361 SMASDGTVAIFISLVNFIIPSKIPLGMQDPKPKGLKAPPAILTWKTVNDKMPNIVILL 420
DB 361 SMASDGTVAIFISLVNFIIPSKIPLGMQDPKPKGLKAPPAILTWKTVNDKMPNIVILL 420
QY 421 GGGFALAKSESGSLSEWLGDKLTPLQHPPTATAVILCLLTAIFTECTSNVATTTFLP 480
DB 421 GGGFALAKSESGSLSEWLGDKLTPLQHPPTATAVILCLLTAIFTECTSNVATTTFLP 480
QY 481 ILASMAQALCLHPLYVYMLPCTLAASLAFMLPVATPPNAIVFSGGLKVDMARAGFLNI 540
DB 481 ILASMAQALCLHPLYVYMLPCTLAASLAFMLPVATPPNAIVFSGGLKVDMARAGFLNI 540
QY 541 IGVLAITLSINSWSIPFKLDTFPPSWAHGNTSQCLNPSNSTVPGGL 587
DB 541 IGVLAITLSINSWSIPFKLDTFPPSWAHGNTSQCLNPSNSTVPGGL 587

RESULT 2
S132_RAT ID -S132_RAT STANDARD; PRT; 587 AA.
AC P70545;
DT 01-NOV-1997 (Rel. 35, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Solute carrier family 13, member 2 (Intestinal sodium/dicarboxylate
DE cotransporter) (Na(+)/dicarboxylate cotransporter 1) (NADC-1).
GN Name=Slc13a2; Synonyms=Nadcl, Sdctl;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar;
RX MEDLINE=97107437; PubMed=8950177; DOI=10.1016/S0167-4781(96)00138-8;
RA Khatiri I.A., Kovacs S.V.B., Forstner J.F.;
RT "Cloning of the cDNA for a rat intestinal Na+/dicarboxylate
RT cotransporter reveals partial sequence homology with a rat intestinal
```

RT mucin.";
 RN [2]
 RA Biophys. Acta 1309:58-62(1996).
 RP CONCEPTUAL TRANSLATION.
 RL Unpublished observations (SEP-2002).
 CC -!- FUNCTION: Cotransport of sodium ions and dicarboxylates such as
 CC succinate and citrate.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -!- SIMILARITY: Belongs to the SLC13A transporter family. NADC
 CC subfamily.
 CC -!- CAUTION: Ref.1 sequence differs from that shown due to the
 CC introduction of a number of frameshifts to maximize the similarity
 CC with orthologs from other species.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 DR EMBL; U51153; AAB97095.1; ALT_FRAME.
 DR RGD; 61920; SLC13a2.
 DR InterPro; IPR001898; Na/sul_sympot.
 DR Pfam; PF00939; Na sulph_symp; 1.
 DR TIGRFAMs; TIGR00785; dase; 1.
 DR PROSITE; PS01271; NA SULFATE; 1.
 KW Sodium transport; Symport; Transmembrane; Transport.
 FT TRANSMEM 13 73 Potential.
 FT TRANSMEM 53 73
 FT TRANSMEM 53 73
 FT TRANSMEM 85 105 Potential.
 FT TRANSMEM 135 155 Potential.
 FT TRANSMEM 263 283 Potential.
 FT TRANSMEM 329 349 Potential.
 FT TRANSMEM 367 387 Potential.
 FT TRANSMEM 407 427 Potential.
 FT TRANSMEM 445 465 Potential.
 FT TRANSMEM 477 497 Potential.
 FT TRANSMEM 506 526 Potential.
 FT TRANSMEM 535 555 Potential.
 SQ SEQUENCE 587 AA; 64206 MW; 964D80661F1EBE0A CRC64;
 Query Match 96.7%; Score 2956; DB 1; Length 587;
 Best Local Similarity 97.4%; Pred. No. 1.8e-215;
 Matches 573; Conservative 2; Mismatches 11; Indels 2; Gaps 2;
 QY 1 MATCHPALWAYRYLVILVLCPLPLPLVQTKAYCAYSILMALLWCTEALPLAVTA 60
 DB 1 MATCHPALWAYRYLVILVLCPLPLPLVQTKAYCAYSILMALLWCTEALPLAVTV 60
 QY 61 LPPVILFPLMGIMDASEVCIEYFKDTNLFVGLMVATAVEHNLHKKRIALQVLLIGVR 120
 DB 61 LPPVILFPLMGIMDASE-GLEYFKDTNLFVGLMVATAVEHNLHKKRIALQVLLIGVR 119
 QY 121 PALLLLGFWLVTAFSLMISNTATTAMVPVIGHAVLEOLQGSKDVGGNNNPTPELOBE 180
 DB 120 PALLLLGFWLVTAFSLMISNTATTAMVPVIGHAVLEOLQGSKDVGGNNNPTPELOBE 179
 QY 181 CPQKEVTKLDNGQPVSAPEPRKTQTBHHRFSQGLSLCICYSASIGIATLTGTPNLV 240
 DB 180 CPQKEVTKLDNGQPVSAPEPRKTQTBHHRFSQGLSLCICYSASIGIATLTGTPNLV 239
 QY 241 LOGQVNSLPQNGVNFASWFGFAPPTMIILLAWLQVLFVGNFRKNFGEGEE 300
 DB 240 LOGQVNSLPQNGVNFASWFGFAPPTMIILLAWLQVLFVGNFRKNFGEGEE 299
 QY 301 ERKQAAFOVTKQYRLLGPMSPAETKVT-TVLVPLVLLVWLFTRPGFFPGWGTVPANSG 359
 DB 300 ERKQAAFOVTKQYRLLGPMSPAETKVT-TVLVPLVLLVWLFTRPGFFPGWGTVPANSG 359
 QY 360 QSMASDGTVAIFISLWFMFIIPSKIPGLMEDPKPKGLKAPPAILLTKVNDKMPWNIVIL 419

DB 360 QSMPSDGTVAIFISLWFMFIIPSKIPGLMEDPKPKGLKAPPAILLTKVNDKMPWNIVIL 419
 QY 420 LGGGFALAKGSEQSGLEWLGDKLTPLOHIPPSSATAVILCLLIAIFTECTSNVATTLFL 479
 DB 420 LGGGFALAKGSEQSGLEWLGDKLTPLOHIPPSSATAVILCLLIAIFTECTSNVATTLFL 479
 QY 480 PILASMAQAICLHPLYVNLPTCLLAASLAFMLPVATPPNAIVPSFGGLKVSMDARAGFLIN 539
 DB 480 PILASMAQAICLHPLYVNLPTCLLAASLAFMLPVATPPNAIVPSFGGLKVSMDARAGFLIN 539
 QY 540 IIGVLATLTSNSWSIPIFKLDTFPPSWAHSNTSOCLNPSNSTVPGGL 587
 DB 540 IIGVLATLTSNSWSIPIFKLDTFPPSWAHSNTSOCLNPSNSTVPGGL 587
 RESULT 3
 S132_MOUSE
 ID S132_MOUSE STANDARD; PRT; 586 AA.
 AC Q95S88;
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Solute carrier family 13, member 2 (Renal sodium/dicarboxylate
 DE cotransporter) (Na(+)/dicarboxylate cotransporter 1) (NADC-1).
 GN Name=SLC13a2; Synonyms=Nadcl, Sdct1;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=20425391; PubMed=10966927;
 RA Pajor A.M., Sun N.N.;
 RT "Molecular cloning, chromosomal organization and functional
 RT characterization of a sodium/dicarboxylate cotransporter from mouse
 RT kidney.";
 RL Am. J. Physiol. 279:F482-F490(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Colon;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Klausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shermen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grilwood J., Schmutz J., Myers R.M.,
 RA Butterfield V.S.N., Krywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -!- FUNCTION: Cotransport of sodium ions and dicarboxylates such as
 CC succinate and citrate.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -!- TISSUE SPECIFICITY: Highly expressed in kidney and small
 CC intestine. Not detectable in brain, heart, stomach and skeletal
 CC muscle.
 CC -!- SIMILARITY: Belongs to the SLC13A transporter family. NADC
 CC subfamily.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
CC EMBL; AF201903; AAG15426.1; -;
CC EMBL; BC013493; AAH13493.1; -;
CC MGD; MGI:1276558; SLC13A2.
CC InterPro; IPR001898; Na/sul_synthet.
CC Pfam; PF00939; Na_sulph_symp; 1.
CC TIGRFAMs; TIGR00785; dss; 1.
CC PROSITE; PS01271; NA_SULFATE; 1.
CC Sodium transporter; Symport; Transmembrane; Transport.
CC TRANSMEM 13 33 Potential.
CC TRANSMEM 53 73 Potential.
CC TRANSMEM 86 106 Potential.
CC TRANSMEM 215 235 Potential.
CC TRANSMEM 264 284 Potential.
CC TRANSMEM 319 339 Potential.
CC TRANSMEM 366 386 Potential.
CC TRANSMEM 407 427 Potential.
CC TRANSMEM 445 465 Potential.
CC TRANSMEM 478 498 Potential.
CC TRANSMEM 506 526 Potential.
CC TRANSMEM 535 555 Potential.
CC SEQUENCE 586 AA; 64110 MW; 56EAE21532033EE CRC64;

Query Match 92.8%; Score 2837; DB 1; Length 586;
Best Local Similarity 92.1%; Pred. No. 1.9e-206;
Matches 539; Conservative 25; Mismatches 21; Indels 0; Gaps 0;

Qy 1 MATCPALWAYRFLVILVLCPIPLPLVOTKEAYCAYSIILMALLWCTEALPLAVTA 60
Db 1 MATCPALWAYRFLVILVLCPIPLPLVOTKEAYCAYSIILMALLWCTEALPLAVTA 60
Qy 61 LFPVILVPLMGLMINDASEVCEYFKDNTLFLVGLMVAIAVEHNLKRIALQVLLIIGVR 120
Db 61 LFPVILVPLMGLMINDASEVCEYFKDNTLFLVGLMVAIAVEHNLKRIALQVLLIIGVR 120
Qy 121 PALLLGLFVLTAFSLMWSNTATTMMVPIGHAVLEQLOGSKDVGGNNPTFELQEE 180
Db 121 PALLLGLFVLTAFSLMWSNTATTMMVPIGHAVLEQLOGSKDVGGNNPTFELQEE 180
Qy 181 CPQEVTKLDNGQVSPASBPRTOKTOEHHRFSSGLSLCYSASIGGIATLTGTTNLY 240
Db 181 SPQKEETKLDNGQVSPASBPRTOKTOEHHRFSSGLSLCYSASIGGIATLTGTTNLY 240
Qy 241 LOGQVNSLFPONGVNVFASFVGFAPFTMTLLLLAWLQVLPLGVNFRKNFGFGE 300
Db 241 LOGQVNSLFPONGVNVFASFVGFAPFTMTLLLLAWLQVLPLGVNFRKNFGFGE 300
Qy 301 ERKQAFQVITQVRLIGPMSFAEKTVTVLVLLVILWFTRBPFGFPFGWGTVPANEKGQ 360
Db 301 ERKQAFQVITQVRLIGPMSFAEKTVTVLVLLVILWFTRBPFGFPFGWGTVPANEKGQ 360
Qy 361 SMASDGTVAIFISIVMTIISKIPGLMOPDKPKLKPAPAILTWKTDKMPNIVILL 420
Db 361 SMASDGTVAIFISIVMTIISKIPGLMOPDKPKLKPAPAILTWKTDKMPNIVILL 420
Qy 421 GGGFALAKGSEGLSEWLGDKLTPLOHIPPASATAVILCLLIAIFTECTSNVATTTFLP 480
Db 421 GGGFALAKGSEGLSEWLGDKLTPLOHIPPASATAVILCLLIAIFTECTSNVATTTFLP 480
Qy 481 ILASMAQAIICLHPIYVLMVLPCTLAASAFMLPVPATPPNAIVFSGGLKVSMDARAGFLINI 540
Db 481 ILASMAQAIICLHPIYVLMVLPCTLAASAFMLPVPATPPNAIVFSGGLKVSMDARAGFLINI 540
Qy 541 IGVLAITLINSWSIPFIKLDTPFSWAHSNTSOCLNPSNTVPG 585
Db 541 IGVLAITLINSWSIPFIKLDTPFSWAHSNTSOCLNPSNTVPG 585

RESULT 4
S132 HUMAN STANDARD; PRT; 592 AA.
AC Q13183;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Solute carrier family 13, member 2 (Renal sodium/dicarboxylate cotransporter) (Na(+)/dicarboxylate cotransporter 1) (NaDC-1).
DE Name=SLC13A2; Synonym=NADC1, SDC11;
GN Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=96199379; PubMed=8967342;
RA Pajor A.M.;
RT "Molecular cloning and functional expression of a sodium-dicarboxylate cotransporter from human kidney";
RL Am. J. Physiol. 270:F642-F648(1996).
CC -!- FUNCTION: Cotransport of sodium ions and dicarboxylates such as succinate and citrate.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: Belongs to the SLC13A transporter family. NADC subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; U26209; AAA98504.1; -;
CC Genew; HGNC:10917; SLC13A2.
CC MTM; 604148; -;
CC GO; GO:0005887; C:integral to plasma membrane; TAS.
CC GO; GO:0005624; C:membrane fraction; TAS.
CC GO; GO:0015361; F:low affinity sodium:dicarboxylate symporter. . .; TAS.
CC GO; GO:0006810; P:transport; TAS.
CC InterPro; IPR001898; Na/sul_symp.
CC Pfam; PF00939; Na_sulph_symp; 1.
CC TIGRFAMs; TIGR00785; dss; 1.
CC PROSITE; PS01271; NA_SULFATE; 1.
CC Sodium transporter; Symport; Transmembrane; Transport.
CC TRANSMEM 13 33 Potential.
CC TRANSMEM 53 73 Potential.
CC TRANSMEM 86 106 Potential.
CC TRANSMEM 114 134 Potential.
CC TRANSMEM 221 241 Potential.
CC TRANSMEM 274 294 Potential.
CC TRANSMEM 324 344 Potential.
CC TRANSMEM 371 391 Potential.
CC TRANSMEM 450 470 Potential.
CC TRANSMEM 482 502 Potential.
CC TRANSMEM 511 531 Potential.
CC TRANSMEM 545 565 Potential.
CC SEQUENCE 592 AA; 64410 MW; 41137D6621A0872A CRC64;

Query Match 78.2%; Score 2392; DB 1; Length 592;
Best Local Similarity 76.9%; Pred. No. 1e-172;
Matches 454; Conservative 63; Mismatches 67; Indels 6; Gaps 4;

Qy 1 MATCPALWAYRFLVILVLCPIPLPLVOTKEAYCAYSIILMALLWCTEALPLAVTA 60
Db 1 MATCPALWAYRFLVILVLCPIPLPLVOTKEAYCAYSIILMALLWCTEALPLAVTA 60
Qy 61 LFPVILVPLMGLMINDASEVCEYFKDNTLFLVGLMVAIAVEHNLKRIALQVLLIIGVR 120

OS Didelphis marsupialis virginiana (North American opossum).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Metatheria; Didelphimorphia; Didelphidae; Didelphis.
 OX NCBI_TaxID=9267;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed=14973148; DOI=10.1152/ajpcell.00061.2003;
 RA Aruga S., Pajor A.M., Nakamura K., Liu L., Moe O.W., Preisig P.A.,
 RA Alpern R.J.;
 RT "ORP cells express the Na-dicarboxylate cotransporter NaDC-1.";
 RL Am. J. Physiol. Cell Physiol. 287:C64-C72(2004).
 DR EMBL; AY186579; AAC27449.1; -;
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005215; F:transporter activity; IEA.
 DR GO; GO:0006814; P:sodium ion transport; IEA.
 DR InterPro; IPR001898; Na/sul_symport.
 DR Pfam; PF00939; Na sulph sym; 1.
 DR PROSITE; PS01271; NA_SULFATE; 1.
 SQ SEQUENCE 605 AA; 67327 MW; 79FAA1AB6421BE7 CRC64;
 Query Match 71.2%; Score 2178; DB 2; Length 605;
 Best Local Similarity 70.0%; Pred. No. 1.7e-156;
 Matches 421; Conservative 67; Mismatches 93; Indels 20; Gaps 5;
 Qy 1 MATCPALWAYRFLYLVLCPLIFLLPLPLIVQTEKAYCAYSIILMALLWCTEALPLAVTA 60
 Db 1 MVNLWRLWAYRKYLSILPILLPLPLVPTKEAKCAYSIILMALLWCTETPLAVTA 60
 Qy 61 LPFVLPLMGIMDASEVCITYKDTNLFVGGMLVAIVAHVHNLHRIALQVLLIIGVR 120
 Db 61 FLPIFFPMGIMDASEVSEIYKDTNLFVGGMLVAIVAHVHNLHRIALQVLLIIGVR 120
 Qy 121 PALLLGLMVLVAFLSWISNTATTMMVPIGHAVLEQLQ--GSKDVEGNNPTELO 178
 Db 121 PAFILGMVTVAFLSWISNTATTMMVPIAHAVLEQLHKGPEEKDTGHWNVISFELQ 180
 Qy 179 EE--CPQKVTKL----DNGQPVSA-----SEPRTKTQBHRFSGQLSLCIC 221
 Db 181 EPHGNPKPEPSLREKENSVPPTSMPEYKEEKEEKEKEKEKEKEKEKEKEKEKEKE 240
 Qy 222 YSASIGIATLTGTPNLVLQGVNSLPFGNNVNFASWFGFAPPTMIILLLLAWLWLQ 281
 Db 241 YASIGIGIATLTGTPNLVLQGVNSLPFGNNVNFASWFGFAPPTMIILLLLAWLWLQ 300
 Qy 282 VFLGVNFRNFGGEGEERKQAFQVKTQYELGLGMSFAKTVTLFVLLVLMFTR 341
 Db 301 ILFLGFNFRNFGGRKSQSKERAAQVQVQTEHKKLGMPSFAETAVTFLFLLVLMFTR 360
 Qy 342 EPGFFPGMGDTVFANEGQSNMADGTVAFISLVMTIPSKIPLQMDPKPKGLKAPPA 401
 Db 361 EPGFFPGMGVNFSGNESGMSVSDGTVAFISILFIVPSLVPFSQO--QGSQKFRAPPA 419
 Qy 402 ILTWKTVNDKPMNVILLGGGFPALAKGSGSGSEWLGBDKLTPLOHIPSATAVILCLL 461
 Db 420 LLDWNTVNMKPMNVILLGGGFPALAKGSEVSGLSMWLGNKLTPLQIPSPATAFILCLL 479
 Qy 462 IATFECTSNVATTTFLPILASNAQAICLHPVLMPLCTLAASLAFMLVATPPNAIVF 521
 Db 480 VATTECANVATTTFLPILASNAQAICLNPVLMPLCTLSASLAFMLVATPPNAIVF 539
 Qy 522 SFGGLKVSMDARAGFLNIIGVLAITLSINSWSIPIFKLDTFFPSWAHSNTSQCLNPSNS 581
 Db 540 SYGOLKVIDMAKTGFLNIIGVLTITLAINTWSYPIFQLDQPTWQAQINSTSQVNGQGN 599
 Qy 582 T 582
 Db 600 T 600
 RESULT 7
 OS 057661
 ID 057661
 AC 057661; PRELIMINARY; PRT; 622 AA.

DT 01-JUN-1998 (TrEMBLrel. 06, Created)
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Intestinal sodium/lithium-dependent dicarboxylate transporter
 DE (NA(+)/dicarboxylate cotransporter).
 DE Names=NADC-2;
 GN Xenopus laevis (African clawed frog).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97423461; PubMed=9277403;
 RA Bai L., Pajor A.M.;
 RT "Expression cloning of NaDC-2, an intestinal Na(+)- or Li(+)-dependent
 RT dicarboxylate transporter.";
 RL Am. J. Physiol. 273:G267-G274(1997).
 CC -I- FUNCTION: TRANSPORTS DI- AND TRICARBOXYLATES, INCLUDING SUCCINATE,
 CC CITRATE, AND GLUTARATE, BUT EXCLUDES THE MONOCARBOXYLATE LACTATE.
 CC -I- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -I- TISSUE SPECIFICITY: Expressed only in the intestine.
 DR EMBL; U87318; AAB97879.1; -;
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0015293; F:symporter activity; IEA.
 DR GO; GO:0005215; F:transporter activity; IEA.
 DR GO; GO:0006814; P:sodium ion transport; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR001898; Na/sul_symport.
 DR Pfam; PF00939; Na sulph sym; 1.
 DR PROSITE; PS01271; NA_SULFATE; FALSE NEG.
 KW Sodium transport; Symport; Transmembrane; Transport.
 FT TRANSMEM 13 33 POTENTIAL.
 FT TRANSMEM 53 73 POTENTIAL.
 FT TRANSMEM 114 134 POTENTIAL.
 FT TRANSMEM 259 279 POTENTIAL.
 FT TRANSMEM 299 319 POTENTIAL.
 FT TRANSMEM 366 386 POTENTIAL.
 FT TRANSMEM 411 431 POTENTIAL.
 FT TRANSMEM 456 476 POTENTIAL.
 FT TRANSMEM 494 514 POTENTIAL.
 FT TRANSMEM 542 562 POTENTIAL.
 FT TRANSMEM 584 604 POTENTIAL.
 FT DOMAIN 123 126 POLY-L-LEU.
 FT CARBOHYD 617 617 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 622 AA; 68550 MW; D2SC06BB806B0098 CRC64;
 Query Match 64.5%; Score 1971.5; DB 2; Length 622;
 Best Local Similarity 63.2%; Pred. No. 8e-141;
 Matches 386; Conservative 87; Mismatches 85; Indels 53; Gaps 10;
 Qy 10 AYRFYLVLCPLIFLLPLPLIVQTEKAYCAYSIILMALLWCTEALPLAVTALPVLPL 69
 Db 10 ANRNVFIPLVPLPLPLVPLVPTKEASGFGVIVNALFWCTEALPLAVTALPVLPL 69
 Qy 70 MGIMDASEVCIEYKDTNLFVGGMLVAIVAHVHNLHRIALQVLLIIGVRPALLLGLM 129
 Db 70 MGIMDSTAVCSQKDTNMLFVGGMLVAIVAHVHNLHRIALQVLLIIGVRPALLLGLM 129
 Qy 130 LVTAFLSMWISNTATTMMVPIGHAVLEQLQSK----KXVEGGNN----- 171
 Db 130 VVTAFLSMWISNTATTMMVPIAQAVMEQLHSSEGVDERVEGNSNTQKNVGMENDMYE 189
 Qy 172 -----NPTFLEQEC-----POKEVTKLDNGQPVSA-----SEPRTKT 206
 Db 190 SYMPSGKMAIDNTYATENEGFEIQEKSTKDPSPSKQ--EKQSIGPIVTEPEDEKQTEK 248
 Qy 207 Q--EHRFSGQLSLCICYSASIGGATLTGTPNLVLQGVNSLPFGNNVNFASWFGF 264
 Db 249 QEKHLKICKGMSLCVCYSASIGGATLTGTPNLVNMKQMDLFPENNININFAWFGF 308
 Qy 265 APFTMIILLLLAWLQVLLVGLVNFNFRKNGFGEGES--ERKQAAFOVIKTQVRLGLPM 322

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2].
RP SEQUENCE FROM N.A.
RC TISSUE=Whole;
RA Strausberg R.;
RL Submitted (DSC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC063947; AAH63947.1; --
DR ZFIN; ZDB-GENE-040426-2389; zgc:77607.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006814; P:sodium ion transport; IEA.
DR InterPro; IPR001898; Na_sulph_symp; 1.
DR Pfam; PF00939; Na_sulph_symp; 1.
SQ SEQUENCE 613 AA; 67653 MW; 8A4B42760D510C6B CRC64;

Query Match 61.1%; Score 1867; DB 2; Length 613;
Best Local Similarity 59.4%; Pred. No. 6.5e-133;
Matches 354; Conservative 101; Mismatches 103; Indels 38; Gaps 9;

Qy 8 LWAYRFLVILCLPILPLPLIVQTKYKAYCAYSIILMALLMCTEALPLAVTALPVLVLF 67
Db 13 LLLHNRNVLILCVPLILPLVLPVTPPEARCGFAIILMALFWCTECPMLAITALLPVVLF 72

Qy 68 PLMGIMDASEVCIEYKDTNLLFVGLMVAIVAEHNLHKLRIALQVLLIIGVPPALLLG 127
Db 73 PMGIMSEGVQVLYKDTNMLFVGLLVAIVAEHNLHKLRIALSVLLVGVPPALLMIG 132

Qy 128 FMLVTAFLSMWISNTATTAMVPIGHAVLEQLQGS-----KDVGGNNNPTFELQE--- 179
Db 133 FMIVTAFLSMWISNTATTAMVPIGHAVLEQLQGS-----KDVGGNNNPTFELQE--- 192

Qy 180 ECP-----QKEVTKLNGQPVSAPEPTQKQE--HHRFSQGLSLCICYSASIGGIAT 231
Db 193 KQPLDNTQGEKPNNAVVEGINALSE-RRRKAREAKYLRFLKGMSLSVCYASIGGTAT 251

Qy 232 LTGTTPNLVLQGVNSLFPONGNVNFAFWGFAFPTMIILLALLMVLQVFLGVNFRK 291
Db 252 LTGTTPNLILKGQWDEIFPDNDVINFAFWGFAFPTMVLVLSWLWLOQVILGFNFQK 311

Qy 292 NFGFGEERKQAAFOVIKTYQVRLGLGPMGFAEKTVTVLVFLVLLVLMFTREPFGFWGD 351
Db 312 SFGCGTKNEGDKO-AYKVMKNEYKKGPMGFAEGAVLVIPLVILWFTREPFGFWGAT 370

Qy 352 TVFANEVQSGWASDGTVAIFSLVWFIIPSKIPGLMQ-----DPKPK 393
Db 371 ELF-KNKGQ-YVTDDGTVAIFSTLFFVIRSDVFLCSIKYEQDEEAVGEQGVKEKK 428

Qy 394 GLKAPPAILTAKVNDKMPNIVILLGGGFALAKGSEQSGLEWLDKTLPLQHIIPPSA 453
Db 429 RLKGTFTLNLKVVHMERMPNIVILLGGGFALAKGSEESGLSVWLGQSPLPSIPPPA 488

Qy 454 TAVILCLLIAIFECTSNVATTLFLPILASMAQAICLHPLYVLMFTCTLAASLAFMLPVA 513
Db 489 ISIIICLLVGTFTCSNTATTLFLPILASMAATTIGLHPLYVLMFTCTISASLAFMLPVA 548

Qy 514 TPNAIVFSGGLKVSMDARAGFIANTIGVLAITLSINWSIPIFKLDTFPPWAHS 569
Db 549 TPNATAFSYGNLKVLDMAKAGFILNIGLICINLGINLTGWMAMFKLDTFPPSWANN 604

RESULT 10
Q803K7
ID Q803K7 PRELIMINARY; PRT; 613 AA.

AC Q803K7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to solute carrier family 13, member 2.
GN ORFNames=zgc:55601;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1].
RP SEQUENCE FROM N.A.
RC STRAIN=AB; TISSUE=Whole body;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2].
RP SEQUENCE FROM N.A.
RC STRAIN=AB; TISSUE=Whole body;
RA Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC044437; AAH44437.1; --
DR ZFIN; ZDB-GENE-040426-2803; zgc:55601.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006814; P:sodium ion transport; IEA.
DR InterPro; IPR001898; Na_sulph_symp; 1.
DR Pfam; PF00939; Na_sulph_symp; 1.
SQ SEQUENCE 613 AA; 67594 MW; F6B73FE343F288C8 CRC64;

Query Match 60.9%; Score 1862; DB 2; Length 613;
Best Local Similarity 59.4%; Pred. No. 1.6e-132;
Matches 354; Conservative 100; Mismatches 104; Indels 38; Gaps 9;

Qy 8 LWAYRFLVILCLPILPLPLIVQTKYKAYCAYSIILMALLMCTEALPLAVTALPVLVLF 67
Db 13 LLLHNRNVLILCVPLILPLVLPVTPPEARCGFAIILMALFWCTECPMLAITALLPVVLF 72

Qy 68 PLMGIMDASEVCIEYKDTNLLFVGLMVAIVAEHNLHKLRIALQVLLIIGVPPALLLG 127
Db 73 PMGIMSEGVQVLYKDTNMLFVGLLVAIVAEHNLHKLRIALSVLLVGVPPALLMIG 132

Qy 128 FMLVTAFLSMWISNTATTAMVPIGHAVLEQLQGS-----KDVGGNNNPTFELQE--- 179
Db 133 FMIVTAFLSMWISNTATTAMVPIGHAVLEQLQGS-----KDVGGNNNPTFELQE--- 192

Qy 180 ECP-----QKEVTKLNGQPVSAPEPTQKQE--HHRFSQGLSLCICYSASIGGIAT 231
Db 193 KQPLDNTQGEKPNNAVVEGINALSE-RRRKAREAKYLRFLKGMSLSVCYASIGGTAT 251

Qy 232 LTGTTPNLVLQGVNSLFPONGNVNFAFWGFAFPTMIILLALLMVLQVFLGVNFRK 291
Db 252 LTGTTPNLILKGQWDEIFPDNDVINFAFWGFAFPTMVLVLSWLWLOQVILGFNFQK 311

RL	Biochem. Biophys. Res. Commun. 299:465-471 (2002).	RL	Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AV151633; AAN86530.1; -	DR	EMBL; AK172785; BAD18766.1; -
DR	Gene; HGNC:23089; SL13A5.	DR	GO; GO:0016020; C-membrane; IEA.
DR	GO; GO:0016020; C-membrane; IEA.	DR	GO; GO:0005215; F-transporter activity; IEA.
DR	GO; GO:0005215; F-transporter activity; IEA.	DR	GO; GO:0006814; P-sodium ion transport; IEA.
DR	InterPro; IPR001898; Na/sul symport.	DR	InterPro; IPR001898; Na/sul symport.
DR	Pfam; PF00939; Na sulph sym; 1.	DR	Pfam; PF00939; Na sulph sym; 1.
DR	PROSITE; PS01271; NA_SULFATE; 1.	DR	PROSITE; PS01271; NA_SULFATE; 1.
SQ	SEQUENCE 568 AA; 63062 MW; B8995E56618DECCB CRC64;	SQ	SEQUENCE 568 AA; 63060 MW; 2535BEEF045C698DB CRC64;
Query Match 52.6%; Score 1624; DB 2; Length 568;			
Best Local Similarity 53.1%; Pred. No. 1.6e-114;			
Matches 309; Conservative 108; Mismatches 132; Indels 32; Gaps 6;			
QY	1 MATCWPAALWAYRYLVLCPIPLFLLPLIVQTKAYCAYSIILMALLWCTEALPLAVTA 60	QY	1 MATCWPAALWAYRYLVLCPIPLFLLPLIVQTKAYCAYSIILMALLWCTEALPLAVTA 60
DB	1 MASALSYSVKFKSFVILFVTPLLLPLVILMPAKFVRCAYVILMAIYMCVEIPLAVTS 60	DB	1 MASALSYSVKFKSFVILFVTPLLLPLVILMPAKFVRCAYVILMAIYMCVEIPLAVTS 60
QY	61 LFPVILFPLMGINDASVCEIYFKDTNLPVGLMVAIAVEHNLKRIALQVLLIIGVR 120	QY	61 LFPVILFPLMGINDASVCEIYFKDTNLPVGLMVAIAVEHNLKRIALQVLLIIGVR 120
DB	61 LMPVLLFPLFQILDSSQVCVQYMKDTNMLFLGLIVAVAVERNLHRIALRTLLWVGAK 120	DB	61 LMPVLLFPLFQILDSSQVCVQYMKDTNMLFLGLIVAVAVERNLHRIALRTLLWVGAK 120
QY	121 PALLLGFMLVTAFLSMWISNTATTAMVPVIGHAVLEQLQSKKQVGGNNNPTFLOEE 180	QY	121 PALLLGFMLVTAFLSMWISNTATTAMVPVIGHAVLEQLQSKKQVGGNNNPTFLOEE 180
DB	121 PARMLGFMGVTTALLSMWISNTATTAMVPVIVEAILQMEATSAATEAG----LELVOK 175	DB	121 PARMLGFMGVTTALLSMWISNTATTAMVPVIVEAILQMEATSAATEAG----LELVOK 175
QY	181 CPQKEVTKLDNGQPVASPSPTQKQEHHRFSQGLSLCICYSASIGGIATLTGTTPNLV 240	QY	181 CPQKEVTKLDNGQPVASPSPTQKQEHHRFSQGLSLCICYSASIGGIATLTGTTPNLV 240
DB	176 GKAKE---LPGSQVIFEGPTLGQOQDEQERKRLCKAMTLCICYAASIGGTATLTGTPNVV 232	DB	176 GKAKE---LPGSQVIFEGPTLGQOQDEQERKRLCKAMTLCICYAASIGGTATLTGTPNVV 232
QY	241 LOGQVNSLPFGNGVNVFASWFGFAPPTMIILLLAWLWQLVFLGVNFRKNGFGEGEE 300	QY	241 LOGQVNSLPFGNGVNVFASWFGFAPPTMIILLLAWLWQLVFLGVNFRKNGFGEGEE 300
DB	233 LLGQNNELFDSKDLNVNFASFAPFNNMLVLLFARLMLQFVYMRFNFKSGCGLESK 292	DB	233 LLGQNNELFDSKDLNVNFASFAPFNNMLVLLFARLMLQFVYMRFNFKSGCGLESK 292
QY	301 ERKQAAFOVTKQYRLGPGMSFAEKTIVTLFVLLVLMFTREPGFPFGW-----GDTV 353	QY	301 ERKQAAFOVTKQYRLGPGMSFAEKTIVTLFVLLVLMFTREPGFPFGW-----GDTV 353
DB	293 KNEKAALKVLEQYRKLGPLSFAEINVLICFFLLVILVFSRDPGFMGMLTVAWVEGETK 352	DB	293 KNEKAALKVLEQYRKLGPLSFAEINVLICFFLLVILVFSRDPGFMGMLTVAWVEGETK 352
QY	354 FANEKGOSMASDGTVAIFISLVMIPIPSKIPGL-----MQDPKPKGLKAPPAITWTKT 407	QY	354 FANEKGOSMASDGTVAIFISLVMIPIPSKIPGL-----MQDPKPKGLKAPPAITWTKT 407
DB	353 Y-----VSDATVAIFVATLTFIVPSQKPNFRSQTEERKTP---FYPPPLLDWKV 401	DB	353 Y-----VSDATVAIFVATLTFIVPSQKPNFRSQTEERKTP---FYPPPLLDWKV 401
QY	408 VNDKMPNVIILGGGFALAKGSEOSGLSEWLGDKLTPLOHIPPSTAVTILCLLIAIFTE 467	QY	408 VNDKMPNVIILGGGFALAKGSEOSGLSEWLGDKLTPLOHIPPSTAVTILCLLIAIFTE 467
DB	402 TQEKVPWGVILLGGGFALAKGSEASGLSVWVGKQMEPLHAVPPAAITLILSLVAVFTE 461	DB	402 TQEKVPWGVILLGGGFALAKGSEASGLSVWVGKQMEPLHAVPPAAITLILSLVAVFTE 461
QY	468 CTSNVATTTFLPILASMAQAICLHPLYVMLPCTLAASLAFMLPVATPPNAIVFSFGGLK 527	QY	468 CTSNVATTTFLPILASMAQAICLHPLYVMLPCTLAASLAFMLPVATPPNAIVFSFGGLK 527
DB	462 CTSNVATTTFLPILASMSRSIGLNPLYIMLPCTLSASFAPMLPVATPPNAIVFTYGHKL 521	DB	462 CTSNVATTTFLPILASMSRSIGLNPLYIMLPCTLSASFAPMLPVATPPNAIVFTYGHKL 521
QY	528 VSDMARAGFLNLIIGVLAITLSINSWSIPIFKLDTFSPSWAH 568	QY	528 VSDMARAGFLNLIIGVLAITLSINSWSIPIFKLDTFSPSWAH 568
DB	522 VADMVKTGVIMNIIGVFCVFLAVNTWGRAIFDLDFPDWVAN 562	DB	522 VADMVKTGVIMNIIGVFCVFLAVNTWGRAIFDLDFPDWVAN 562
RESULT 14			
Q67BT3	PRELIMINARY; PRT; 572 AA.	Q67BT3	PRELIMINARY; PRT; 572 AA.
AC	Q67BT3; (TrEMBLrel. 28, Created)	AC	Q67BT3; (TrEMBLrel. 28, Created)
DT	25-OCT-2004 (TrEMBLrel. 28, Last sequence update)	DT	25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT	25-OCT-2004 (TrEMBLrel. 28, Last annotation update)	DT	25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE	Sodium-coupled citrate transporter.	DE	Sodium-coupled citrate transporter.
GN	Names=Nact;	GN	Names=Nact;
OS	Mus musculus (Mouse).	OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;		NCBI_TaxID=10090;	
RN	[1]	RN	[1]
RP	SEQUENCE FROM N.A.	RP	SEQUENCE FROM N.A.
RA	Inoue K., Fei Y.-J., Zhuang L., Gopal B., Miyauchi S., Ganapathy V.;	RA	Inoue K., Fei Y.-J., Zhuang L., Gopal B., Miyauchi S., Ganapathy V.;
RT	"functional features and genomic organization of mouse NaCT, a sodium-	RT	"functional features and genomic organization of mouse NaCT, a sodium-
RT	coupled transporter for citric acid cycle intermediates."	RT	coupled transporter for citric acid cycle intermediates."

THIS PAGE IS BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2005, 08:28:14 ; Search time 84.655 Seconds
(without alignments)
2741.199 Million cell updates/sec

Title: US-10-017-479A-5
Perfect score: 3094
Sequence: 1 MAALALAKVWSARRLLVL.....HAANVTALPPALTNTVTQL 600

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : A_Geneseq_16Dec04.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003s.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2741	88.6	602	6	ABR40097 Human sod
2	2741	88.6	602	6	ABR40101 Human tra
3	2741	88.6	602	6	ABG75837 Transport
4	2326.5	75.2	520	7	ADM05177 Human pro
5	2231.5	72.1	533	8	ABM84303 Human dia
6	1412.5	45.7	616	5	ABU65064 Human NOV
7	1412.5	45.7	616	5	ABU65063 Human NOV
8	1412.5	45.7	616	7	ADK51048 Human NOV
9	1412.5	45.7	616	8	ADH42447 Novel hum
10	1412.5	45.7	616	8	ADN61777 Human nov
11	1412.5	45.7	616	8	ADN61779 Human nov
12	1411.5	45.6	568	6	ABR57023 Human TCH
13	1411.5	45.6	568	7	AAE38764 Human 696
14	1411.5	45.6	568	7	ADK51052 Human NOV
15	1411.5	45.6	568	8	ADH42441 Novel hum
16	1411.5	45.6	568	8	ADP64793 Human Na+
17	1411.5	45.6	576	8	ADH42443 Novel hum
18	1405.5	45.4	568	7	ADE07994 Novel pro
19	1405.5	45.4	619	8	ADG16995 African c
20	1402.5	45.3	568	6	ABG75835 Transport
21	1400.5	45.3	568	5	ABU65062 Human NOV
22	1400.5	45.3	568	7	ADK51050 Human NOV
23	1400.5	45.3	568	8	ADH42445 Novel hum
24	1400.5	45.3	568	8	ADN61775 Human nov
25	1399	45.2	581	8	ADP64799 Zebrafish

99 737 23.8 510 4 AAG90000 C glutami
100 699.5 22.6 432 6 ABM71452 Staphyloc

ALIGNMENTS

RESULT 1
ABB82952
ID ABB82952 standard; protein; 602 AA.
XX
AC ABB82952;
XX
DT 14-APR-2003 (first entry)
XX
DE Human SLC13A related protein (GenBank Identifier No. GI#13653602).
XX
KW SLC13A; p53; sodium-sulfate cotransporter 2; cytostatic; cancer;
KW transmembrane protein; human.
XX
OS Homo sapiens.
XX
PN WC020298468-A1.
XX
PD 12-DEC-2002.
XX
PF 03-JUN-2002; 2002WO-US017460.
XX
PR 05-JUN-2001; 2001US-0296076P.
PR 10-OCT-2001; 2001US-0328605P.
PR 15-FEB-2002; 2002US-0357253P.
PR 01-MAR-2002; 2002US-0361196P.
XX
XX (EXEL-) EXELIXIS INC.
XX
XX Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
XX Lioubin MN;
XX WFI; 2003-167297/16.
XX
XX Identifying candidate p53 pathway modulating agent for diagnosing or
XX treating cancer comprises detecting test agent-biased activity of an
XX assay system comprising purified Sodium sulfate cotransporter 2 (SLC13A)
XX polypeptide.
XX
XX Claim 13; Page 64-67; 69pp; English.
XX

CC The invention relates to identifying a candidate p53 pathway modulating
CC agent. The method involves assaying purified Sodium-sulfate cotransporter
CC 2 (SLC13A) polypeptide or nucleic acid or a functionally active fragment
CC or derivative with a test agent. The methods are useful for identifying a
CC candidate p53 pathway modulating agent, modulating a p53 pathway of a
CC cell, or a mammalian cell and for diagnosing a disease in a patient. The
CC disease is breast, colon, lung or ovary cancer having greater than 25%
CC expression level. The method is useful for manufacturing a medicament for
CC diagnosing or treating breast, colon, lung or ovary cancer. Sequences
CC ABB82949-952 represent polypeptide sequences related to the human SLC13A
CC protein
XX
SQ Sequence 602 AA;
Query Match 88.6%; Score 2741; DB 6; Length 602;
Best Local Similarity 87.4%; Pred. No. 1.1e-278;
Matches 526; Conservative 41; Mismatches 33; Indels 2; Gaps 2;
1 MAALAAKAKVWSARRLLVLLVPLALLPILFALPPKRGCLYVILLMAVYVCTEALPLS 60
1 MAALAAKAKVWSARRLLVLLVPLALLPILFALPPKRGCLYVILLMAVYVCTEALPLS 60
61 VTALLPILPFPMGILPSSKVCPOYFDTNPLFSLGILMASAEERNLHRIALKVILMV 120
61 VTALLPILPFPMGILPSSKVCPOYFDTNPLFSLGILMASAEERNLHRIALKVILMV 120

QY 121 GVQPARLILGMVTTSTFLSMMLNSTASTAMMLPIASAILKSLFGQDTRKDLPRGEDST 180
DB 121 GVQPARLILGMVTTSTFLSMMLNSTASTAMMLPIANAILKSLFGQKEVRKDPQSSEENT 180
QY 181 AAVRNGRLTVPTQMQLASSEG-GHAEDVEAPLELDDDS-KEEHRNRNIWGLFISIPY 238
DB 181 AAVRNGRLTVPTQMQLASTEAKDHPGETEVPLDLPADSRKEDERYRRNIWGLFISIPY 240
QY 239 SASIGGTATLTGTAPNLILLGOLKSFPPQCDVNVFGSWFIFAPFLMLLFLVGLWLSFL 298
DB 241 SASIGGTATLTGTAPNLILLGOLKSFPPQCDVNVFGSWFIFAPFLMLLFLVGLWLSFL 300
QY 299 YGMSWRGWRKNSKLDQVAEDKAKAVIQEESQNLGPIKFAEQAVFILLFCLFAILLFSRD 358
DB 301 YGGLSFRGWRKNSKIRTNADBARAVIREEQNLGPIKFAEQAVFILLFCLFAILLFTRD 360
QY 359 PKFIPGWSLFPAGFVSDAVTGVAVITLIFPPPSQKPSLKWDFDKAPNSETETPLLSWK 418
DB 361 PKFIPGWSLFPAGFVSDAVTGVAVITLIFPPPSQKPSLKWDFDKAPNSETETPLLSWK 420
QY 419 AQETVPMNIIILLGGGFAMKCESGLSAGTGGQLHPLHLEHVPPLAVLLITVIAFFTE 478
DB 421 AQETVPMNIIILLGGGFAMKCESGLSAGTGGQLHPLHLEHVPPLAVLLITVIAFFTE 480
QY 479 FASNTATIIIFLPVLAELAIRLHVHPLVLMIPGTVSCSYAFMLPVSTPNSIAFSGHLL 538
DB 481 FASNTATIIIFLPVLAELAIRLHVHPLVLMIPGTVSCSYAFMLPVSTPNSIAFSGHLL 540
QY 539 VKDMVRTGLLMLMGVLLLSLAMNTWAQAI FOLGTFPPDWANTHAANVTALPPALNTNTVQ 598
DB 541 VKDMVRTGLLMLMGVLLLSLAMNTWAQAI FOLGTFPPDWANTHAANVTALPPALNTNTVQ 600
QY 599 TL 600
DB 601 TL 602
RESULT 2
ABR40101
ID ABR40101 standard; protein; 602 AA.
XX
AC ABR40101;
XX
DT 24-JUL-2003 (first entry)
XX
DE Human renal sodium dicarboxylate co-transporter NADC3.
XX
KW Anti-diabetic; anorectic; sodium dicarboxylate cotransporter; human;
KW sodium tricarboxylate cotransporter; HepNadC; hepatocyte; HepG2;
KW diabetes; obesity; lipid metabolism; aging; NADC3.
XX
OS Homo sapiens.
XX
PN WO2003029465-A1.
XX
PD 10-APR-2003.
XX
PF 27-SEP-2002; 2002WO-JP010038.
XX
PR 28-SEP-2001; 2001JP-00299433.
PR 28-AUG-2002; 2002JP-00249016.
XX
XX (SAKA) OTSUKA PHARM CO LTD.
XX Kanemoto N, Omori Y, Sugano S, Obuchi Y;
XX WFI; 2003-354728/33.
XX
XX Sodium di- or tricarboxylate cotransporter gene (HepNadC) for treatment
XX and prevention of diabetes, obesity, for improving lipid metabolism and
XX aging.
XX
XX Example 1; Fig 1; 75pp; Japanese.

XX The present invention relates to human sodium di- or tricarboxylate
CC cotransporter protein (HepNADC; ABR40097). HepNADC gene is expressed in
CC human hepatocyte cancer-origin cells HepG2. Compounds which control the
CC expression of the HepNADC gene and activity of the HepNADC polypeptide
CC may be useful for treating and preventing diabetes, obesity, for
CC improving lipid metabolism and aging. The present sequence is the NADC3
CC protein (GenBank XM_017841), which was used in a sequence alignment with
XX the HepNADC sequence
SQ Sequence 602 AA;
Query Match 88.6%; Score 2741; DB 6; Length 602;
Best Local Similarity 87.4%; Pred. No. 1.1e-278;
Matches 526; Conservative 41; Mismatches 33; Indels 2; Gaps 2;
QY 1 MAALAAAKKWSARRLLVLLVPLALPILFALPPKGRCLYVILLMAVYVCTEALPLS 60
DB 1 MAALAAAKKWSARRLLVLLVPLALPILFALPPKGRCLYVILLMAVYVCTEALPLS 60
QY 61 VTALLPIILFPPMGLIPSKVCPQYFLDTNFTLFLSGLIMASAIERNLHRRIALKVLMLV 120
DB 61 VTALLPIILFPPMGLIPSKVCPQYFLDTNFTLFLSGLIMASAIERNLHRRIALKVLMLV 120
QY 121 GVQPARLLIGMVTTSFSLMWSLNTASTAMMLPIASAILKSLFGQDRTRKPLRGEDST 180
DB 121 GVQPARLLIGMVTTSFSLMWSLNTASTAMMLPIASAILKSLFGQDRTRKPLRGEDST 180
QY 121 GVQPARLLIGMVTTSFSLMWSLNTASTAMMLPIANAILKSLFGQKVRKDPQSESENT 180
DB 121 GVQPARLLIGMVTTSFSLMWSLNTASTAMMLPIANAILKSLFGQKVRKDPQSESENT 180
QY 181 AAVRNGRLTVPTEMQFVLASSG-GBAEDVAPLELPDS-KEREHRRNINWGFLLISIPY 238
DB 181 AAVRNGRLTVPTEMQFVLASSG-GBAEDVAPLELPDS-KEREHRRNINWGFLLISIPY 238
QY 181 AAVRNGRLTVPTEMQFVLASSG-GBAEDVAPLELPDS-KEREHRRNINWGFLLISIPY 240
DB 181 AAVRNGRLTVPTEMQFVLASSG-GBAEDVAPLELPDS-KEREHRRNINWGFLLISIPY 240
QY 239 SASIGGTATLTGAPNLILLGOLKSFPPQCDVNVFSGSWFIFAPPLMLLFLVGLWLSIFL 298
DB 241 SASIGGTATLTGAPNLILLGOLKSFPPQCDVNVFSGSWFIFAPPLMLLFLVGLWLSIFL 300
QY 299 YGGSWRGWRKNSKLOVADKAKAVIOEFQNGPIKFAEQVPIFLCFLFAILLFSD 358
DB 301 YGGLSFRGWRKNSKIRTNADRAVARVIREYQNLGPIKFAEQVPIFLCFLFAILLFTRD 360
QY 359 PKFIPGWSLAPGVPVDAVTGVAITVILFFPPSKPSLKWDFKAPNSETEPLLSWK 418
DB 361 PKFIPGWSLAPGVPVDAVTGVAITVILFFPPSKPSLKWDFKAPNSETEPLLSWK 420
QY 419 AQETVPWNIILLGGGFANAKGCEBSGLSAMTIGGQHPLEHVPPLLAVLLTVIAFTE 478
DB 421 AQETVPWNIILLGGGFANAKGCEBSGLSAMTIGGQHPLEHVPPALAVLLTVIAFTE 480
QY 479 PASNTATIIIFLPVLAELAIRLHVHPLYLMIPGTVCSSYAFMLPVSTPPNSIAFSTGHL 538
DB 481 PASNTATIIIFLPVLAELAIRLHVHPLYLMIPGTVCSSYAFMLPVSTPPNSIAFSTGHL 540
QY 539 VKDMVRTGLLNLGVLISLAMNTWQAIIQLGTFPPDMWANTHAANTVLPALTNVQ 598
DB 541 VKDMVRTGLLNLGVLISLAMNTWQAIIQLGTFPPDMWANTVSNVTPALPTLANDTFR 600
QY 599 TL 600
DB 601 TL 602
RESULT 3
ABG75837
ID ABG75837 standard; protein; 602 AA.
XX AC ABG75837;
XX DT 10-MAY-2003 (first entry)
XX DE Transporters and ion channels protein 19, TRICH-19.
XX KW Human; transporter and ion channel; TRICH; gene therapy;
KW cell proliferative disorder; transport disorder; neurological disorder;

KW muscle disorder; immunological disorder; amyotrophic lateral sclerosis;
KW cystic fibrosis; diabetes; Parkinson's disease; prostate cancer;
KW cardiac disorder; angina; Alzheimer's disease; amnesia; epilepsy;
KW schizophrenia; sickle cell anaemia; infertility; hyperglycaemia;
KW hypoglycaemia; hypercholesterolaemia; stroke; multiple sclerosis;
KW motor neuron disorder; prion disease; metabolic disease;
KW developmental disorder; central nervous system; cardiomyopathy;
KW hypertension; asthma; AIDS; allergy; anaemia; atherosclerosis;
KW atopic dermatitis; diabetes mellitus; osteoarthritis; osteoporosis;
KW rheumatoid arthritis; psoriasis; infection; trauma; hepatitis; cancer;
KW leukemia; lymphoma.
XX Homo sapiens.
OS WO2003016493-A2.
XX 27-FEB-2003.
XX 16-AUG-2002; 2002WO-US026323.
XX 17-AUG-2001; 2001US-0313242P.
PR 21-SEP-2001; 2001US-0324782P.
PR 02-OCT-2001; 2001US-0328184P.
PR 26-OCT-2001; 2001US-0345937P.
PR 01-NOV-2001; 2001US-0335698P.
PR 13-NOV-2001; 2001US-0332804P.
PR 27-NOV-2001; 2001US-0333922P.
PR 26-APR-2002; 2002US-0375637P.
PR 03-MAY-2002; 2002US-0377444P.
PR 11-JUN-2002; 2002US-0388180P.
XX (INCY-) INCYTE GENOMICS INC.
XX Tang YT, Lal PG, Yue H, Baughn MR, Nguyen DB, Yao MG, Greene BD;
PI Borowsky ML, Lee S, Emerling BM, Xu Y, Becha SD, Gorvad AE;
PI Azimzai Y, Yue H, Elliott VS, Lee EA, Yang J, Lehr-Watson PW;
PI Rankumar J, Lee SY, Faris M, Turner C, Furness M, Buchbinder JL;
PI Walia NK, Li JX, Forsythe LJ, Griffin JA, Gietzen KJ, Swarnakar A;
PI Hafalia AUA, Lindquist EA, Jiang X, Jackson AA, Wilson AD, Jin P;
PI Khare R, Marquis JP;
XX WPI; 2003-268319/26.
DR N-PSDB; ABX12022.
XX Novel human transporter and ion channel polypeptides and polynucleotides
for diagnosing, preventing or treating cell proliferative, transport,
neurological, muscle and immunological disorders.
Claim 1; Page 221-222; 253pp; English.
The invention discloses isolated polypeptides chosen from human
transporter and ion channel polypeptides, TRICH 1-26, a biologically
active or immunogenic fragment and the nucleic acids encoding them. Also
disclosed are isolated antibodies raised against the TRICH proteins,
methods for detecting a target polynucleotide in a sample and a
microarray where at least one element is a TRICH polynucleotide. The
proteins are useful for screening for agonists or antagonists, which can
then be used for treating a disease or condition associated with
decreased or overexpression of functional TRICH in a patient, for
screening for a compound that modulates the activity of the polypeptide
or that binds to the polypeptide or as an immunogen for preparing
antibodies. The polynucleotides are useful for screening for compounds
which alter expression of a target polynucleotide or for assessing
toxicity of a test compound. The polypeptides, polynucleotides,
modulators and antibodies are useful for diagnosis, treatment (e.g. gene
therapy) and prevention of cell proliferative, transport, neurological,
muscle and immunological disorders, such as amyotrophic lateral
sclerosis, cystic fibrosis, diabetes, Parkinson's disease, prostate
cancer, cardiac disorders, angina, Alzheimer's disease, amnesia,
epilepsy, schizophrenia, sickle cell anaemia, infertility,
hyperglycaemia, hypoglycaemia, hypercholesterolaemia, stroke, multiple
sclerosis, motor neuron disorder, prion disease, metabolic disease of the
nervous system, developmental disorders of the central nervous system,

CC cardiomyopathy, hypertension, asthma, AIDS, allergies, anaemia.
CC atherosclerosis, atopic dermatitis, diabetes mellitus, osteoarthritis,
CC osteoporosis, rheumatoid arthritis, psoriasis, infections, trauma,
CC hepatitis and cancers, including leukemia and lymphoma. The sequences
CC presented in ABG75819-ABG75844 are the TRICH proteins of the invention
XX
SQ Sequence 602 AA;

Query Match 88.6%; Score 2741; DB 6; Length 602;
Best Local Similarity 87.4%; Pred. No. 1.1e-278;
Matches 526; Conservative 41; Mismatches 33; Indels 2; Gaps 2;

QY 1 MAALALAKVWSARLLVLLVPLALLPILPALPKGRCCLVILLMAVYVCTEALPLS 60
DB 1 MAALAAAKVWSARLLVLLVPLALLPILPALPKGRCCLVILLMAVYVCTEALPLS 60

QY 61 VTALLPIILFPFNGILPSSKVCQYFELDTNLFSLGLIMASAEERNLHRRALKVLMVL 120
DB 61 VTALLPIVLPFNGILPSSKVCQYFELDTNLFSLGLIMASAEERNLHRRALKVLMVL 120

QY 121 GVOPARLILGMVTTFSLSNLSNTASTAMMLPIASAILKSLFGQDTRKDLPRGEDST 180
DB 121 GVOPARLILGMVTTFSLSNLSNTASTAMMLPIANAILKSLFGQEVKRDPSQSEENT 180

QY 181 AAVRGNGLRTVPTMOPFLASSEG-GHAEDVEAPLELPDDDS-KKEEHRNIWKGLISIPY 238
DB 181 AAVRRNGLHTVPTMOPFLASTEAKDHPGETEVFLDLPADSRKDEYRNINWKGLISIPY 240

QY 239 SASIGGTATLTGTAPNLILGLQKSPFPQCDVVNFGSWFIFAPPLMLFLVGLMWISFL 298
DB 241 SASIGGTATLTGTAPNLILGLQKSPFPQCDVVNFGSWFIFAPPLMLFLVGLMWISFL 300

QY 299 YGMSWRGWRKNSKLODVADKAKAVIQEFQNLGPIKFAEQAVRILFCLFAILLPSRD 358
DB 301 YGGLSFRGWRKNSKSEIRTNADRARAVIREEQNLGPIKFAEQAVRILFCLFAILLPSRD 360

QY 359 PKPIPGWASLFPAGFVSDAVTGVAIVTILFFFPQSKPSLKWDFKAPNSETEPILLSWK 418
DB 361 PKPIPGWASLFPAGFVSDAVTGVAIVTILFFFPQSKPSLKWDFKAPNSETEPILLSWK 420

QY 419 AQETVPWNIIILGGGFAMAKGCEESGLSAWIGQLHPLHPLHPLHPLHPLHPLHPLHPL 478
DB 421 AQETVPWNIIILGGGFAMAKGCEESGLSAWIGQLHPLHPLHPLHPLHPLHPLHPLHPL 480

QY 479 FASNTATIIIFLPLVLAELAIRLHVHPLVLMIPGTVSCSYAFMLPVSTPPNSIAFSTGHLL 538
DB 481 FASNTATIIIFLPLVLAELAIRLHVHPLVLMIPGTVSCSYAFMLPVSTPPNSIAFSTGHLL 540

QY 539 VKDMVRTGLLNMVGLVLLSLAMNTWAQAIIFOLGTFDPDMANTHAANVTALPPALTNNVTQ 598
DB 541 VKDMVRTGLLNMVGLVLLSLAMNTWAQTIIFOLGTFDPDMANTHAANVTALPPALTNNVTQ 600

QY 599 TL 600
DB 601 TL 602

RESULT 4
ADM05177
ID ADM05177 standard; protein; 520 AA.
XX
AC ADM05177;
XX
DT 20-MAY-2004 (first entry)
XX
DE Human protein of the invention SEQ ID NO:3862.
XX human; gene therapy; diagnostic marker; pharmaceutical.
XX Homo sapiens.
XX EP1347046-A1.
XX

PD 24-SEP-2003.
XX
PF 12-APR-2002; 2002EP-00008400.
XX
PR 22-MAR-2002; 2002JP-00137785.
XX
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI Isozaki T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuno Y;
XX
XX WPI: 2003-723558/69.
DR N-PSDB; ADM02734.
XX
XX New polynucleotides and polypeptides are useful in gene therapy, for
PT developing a diagnostic marker or medicines for regulating their
PT expression and activity, or as a target of gene therapy.
XX
XX Claim 1; SEQ ID NO 3862; 305pp; English.
XX
XX The invention relates to a novel human polynucleotide and the encoded
XX polypeptide. A polynucleotide of the invention may have a use in gene
XX therapy. An oligonucleotide of the invention ADM06202-ADM06773 is useful
XX as a primer for synthesizing the polynucleotide or as a probe for
XX detecting the polynucleotide. The polynucleotides ADM01316-ADM03758 are
XX useful in gene therapy, for developing a diagnostic marker or medicines
XX for regulating their expression and activity, or as a target of gene
XX therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides
XX are useful as pharmaceutical agents. The present sequence represents a
XX protein sequence of the invention.
SQ Sequence 520 AA;

Query Match 75.2%; Score 2326.5; DB 7; Length 520;
Best Local Similarity 81.4%; Pred. No. 3.7e-235;
Matches 452; Conservative 36; Mismatches 30; Indels 37; Gaps 3;

QY 48 MAVYWCTEALPLSVTALLPIILFPFNGILPSSKVCQYFELDTNLFSLGLIMASAEERN 107
DB 1 MAVYWCTEALPLSVTALLPIILFPFNGILPSSKVCQYFELDTNLFSLGLIMASAEERN 60

QY 108 LHRRIALKVLMVGVOPARLILGMVTTFSLSNLSNTASTAMMLPIASAILKSLFGQD 167
DB 61 LHRRIALKVLMVGVOPARLILGMVTTFSLSNLSNTASTAMMLPIANAILKSLFGQKE 120

QY 168 TRKDLPRGEDSTAAVRGNGLRTVPTMOPFLASSEG-GHAEDVEAPLELPDDDS-KKEEHR 225
DB 121 VRKDPQSEENTAAVRNGLHTVPTMOPFLASTEAKDHPGETEVFLDLPADSRKDEYR 180

QY 226 RNIWKGLISIPYSASIGGTATLTGTAPNLILGLQKSPFPQCDVVNFGSWFIFAPPLML 285
DB 181 RNIWKGLISIPYSASIGGTATLTGTAPNLILGLQKSPFPQCDVVNFGSWFIFAPPLML 240

QY 286 LFLVGLMWISFLYGGMSWRGWRKNSKLODVADKAKAVIQEFQNLGPIKFAEQAVFI 345
DB 241 LFLVGLMWISFLYGGMSWRGWRKNSKSEIRTNADRARAVIREEQNLGPIK- 292

QY 346 LFCLFAILLPSRDPKFIPGWASLFPAGFVSDAVTGVAIVTILFFFPQSKPSLKWDFDKA 405
DB 293 -----FLSDAVTGVAIVTILFFFPQSKPSLKWDFDKA 325

QY 406 PNSETEPILLSWKKAQETVPWNIIILGGGFAMAKGCEESGLSAWIGQLHPLHPLHPL 465
DB 326 PNTETEPILLSWKKAQETVPWNIIILGGGFAMAKGCEESGLSAWIGQLHPLHPLHPL 385

QY 466 VLLITVVIAPFTEFASNTATIIIFLPLVLAELAIRLHVHPLVLMIPGTVSCSYAFMLPVST 525
DB 386 VLLITVVIAPFTEFASNTATIIIFLPLVLAELAIRLHVHPLVLMIPGTVSCSYAFMLPVST 445

QY 526 PPNISAFSTGHLLVAKDMVRTGLLNMVGLVLLSLAMNTWAQAIIFOLGTFDPDMANTHAAN 585
DB 446 PPNISAFASGHLVAKDMVRTGLLNMVGLVLLSLAMNTWAQTIIFOLGTFDPDMANTHAAN 505

QY 586 TALPPALNTNTVQTL 600
 DB 506 TALPPTLANDTPTTL 520

RESULT 5
 ABM84303
 ID ABM84303 standard; protein; 533 AA.
 AC
 XX
 XX
 DT 18-NOV-2004 (first entry)
 DE Human diagnostic and therapeutic pprotein SEQ ID NO:4552.
 DE Human diagnostic and therapeutic polynucleotide; dithp.
 KW gene therapy; human diagnostic and therapeutic polynucleotide; dithp.
 OS Homo sapiens.
 XX
 XX WO2004023973-A2.
 PN 25-MAR-2004.
 PD 12-SEP-2003; 2003WO-US028227.
 PF 12-SEP-2002; 2002US-0410259P.
 PR 12-SEP-2002; 2002US-0410260P.
 XX (INCY-) INCYTE CORP.
 PA Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
 PI Harthshorne TA, Suchorolski MT, Altus CM, Pitts SJ, Blder LV;
 PI Money EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;
 PI Stevens KA, Blanchard JB, Panzer SR, Wang X, Au AP, Gerstin EH;
 PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
 PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES;
 PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
 PI Patuary S, Shi X, Suarez CJ;
 XX WPI: 2004-329368/30.
 DR N-PSDB; ACN42955.
 XX
 XX New diagnostic and therapeutic polynucleotides and polypeptides, useful
 PT in diagnosing a condition, disease or disorder associated with human
 PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
 PT in gene mapping.
 XX
 XX Claim 27; Page: 190pp; English.
 PS
 XX The invention relates to novel diagnostic and therapeutic polynucleotides
 CC selected from one of the 2722 sequences defined in the specification. A
 CC polynucleotide of the invention may have a use in gene therapy. The human
 CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
 CC used to diagnose a particular condition, disease or disorder associated
 CC with human molecules, e.g. cell proliferative disorders,
 CC autoimmune/inflammatory disorder, developmental disorder, endocrine
 CC disorder, neurological disorders, gastrointestinal disorders, or
 CC infections caused by virus, bacteria, fungi or parasite. The dithp
 CC molecules may also be used in genetic mapping, in identifying individuals
 CC from minute biological samples, in detecting single nucleotide
 CC polymorphisms, as molecular weight markers, and for somatic or germ-line
 CC gene therapy. The present sequence represents a dithp protein of the
 CC invention. Note: The sequence data for this patent is not represented in
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at www.wipo.int/pct/en/sequences/listing.htm
 XX

Query Match 72.1%; Score 2231.5; DB 8; Length 533;
 Best Local Similarity 73.8%; Pred. No. 3.8e-225;
 Matches 445; Conservative 39; Mismatches 46; Indels 73; Gaps 5;

QY 1 MAALAAAKKWSARRLLVLLVPLALLPILPALPPKGGRCRLYVILLMAVYCTEALPLS 60
 DB 1 MAALAAAKKWSARRLLVLLVPLALLPILPALPPKGGRCRLYVILLMAVYCTEALPLS 59
 QY 61 VTALLPIILFPFMGILPSSKVCPOVFLDTNLFSLGLIMASAEIRNLHRRITALKVLMV 120
 DB 60 VTALLPIILFPFMGILPSSKVCPOVFLDTNLFSLGLIMASAEIRNLHRRITALKVLMV 119
 QY 121 GVQPARLILGMVTTTSLFSLMWSLNTASTAMMLPIASAILKSLFGQDRKDLPRGEDST 180
 DB 120 GVQPARLILGMVTTTSLFSLMWSLNTASTAMMLPIANAILKSLFGQKEVRKQDSQSEENT 179
 QY 181 AAVRGNGLRTV-PTMQFLASSEG-GHAEDVEAPLELPDD-KEBEHRNINWKGFLISIP 237
 DB 180 GEAVGETAYTLCPRRCSPXPTEAKDHPGETEVLPLDLPADSRKDEYRRNINWKGFLISIP 239
 QY 238 YSASIGGTATLTGTAPNLILGOLKSPFQCDVNVFGSWFIFAPFLMLLFLLVGMLWISF 297
 DB 240 YSASIGGTATLTGTAPNLILGOLKSPFQCDVNVFGSWFIFAPFLMLLFLLVGMLWISF 299
 QY 298 LYGGMSWRGMRKNSKLQDVAEDKAKAVIOBEFQNLGPIKFAEQAVFIFLCLFALLPSR 357
 DB 300 LYGGLSFRGMRKNSKSEIRTNADRARAVIREYQNLGP-----APNTEPELTLTWK 350
 QY 358 DPKFIPGWASLFAFGFVSDAVTGVAIVTILFFFPSPKPSLKWDFDKAPNSETPELTLGWK 417
 DB 338 -----APNTEPELTLTWK 350
 QY 418 KAQETVPWNIIILGGGFAMAKGCEESGLSAGIQLHPLHVPPLAVLITVVIAPFT 477
 DB 351 KAQETVPWNIIILGGGFAMAKGCEESGLSAGIQLHPLHVPPLAVLITVVIAPFT 410
 QY 478 EFASNTATIIIFLPVLAELAIRLHVHPLYLMIPGTVSCSYAFMLPVSTPPNSIAFSTGHL 537
 DB 411 EFASNTATIIIFLPVLAELAIRLHVHPLYLMIPGTVSCSYAFMLPVSTPPNSIAFSTGHL 470
 QY 538 LVKDMVRTGLMNLMLGVLISLAWNTWAQAIIFQLGTPDWMANTHAANYTALPPALTNTV 597
 DB 471 LVKDMVRTGLMNLMLGVLISLAWNTWAQAIIFQLGTPDWMANTHAANYTALPPALTNTV 530
 QY 598 QTL 600
 DB 531 RTL 533

RESULT 6
 ABU65064
 ID ABU65064 standard; protein; 616 AA.
 XX
 AC ABU65064;
 XX
 DT 20-MAY-2003 (first entry)
 DE Human NOV14c protein.
 XX
 KW NOVX; cytostatic; cardiant; antiarteriosclerotic; antiasthmatic; cancer;
 KW hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;
 KW human.
 XX
 OS Homo sapiens.
 XX
 PN WO200272757-A2.
 XX
 PD 19-SEP-2002.
 XX
 PF 08-MAR-2002; 2002WO-US006908.
 XX
 PR 08-MAR-2001; 2001US-0274101P.
 PR 08-MAR-2001; 2001US-0274194P.
 PR 08-MAR-2001; 2001US-0274281P.
 PR 08-MAR-2001; 2001US-0274322P.
 PR 09-MAR-2001; 2001US-0274849P.
 PR 12-MAR-2001; 2001US-0275235P.

PR 13-MAR-2001; 2001US-0275578P.
PR 13-MAR-2001; 2001US-0275579P.
PR 13-MAR-2001; 2001US-0275601P.
PR 14-MAR-2001; 2001US-0276000P.
PR 16-MAR-2001; 2001US-0276776P.
PR 19-MAR-2001; 2001US-0276994P.
PR 20-MAR-2001; 2001US-0277329P.
PR 20-MAR-2001; 2001US-0277321P.
PR 20-MAR-2001; 2001US-0277327P.
PR 21-MAR-2001; 2001US-0277791P.
PR 22-MAR-2001; 2001US-0277833P.
PR 23-MAR-2001; 2001US-0278152P.
PR 26-MAR-2001; 2001US-0278894P.
PR 27-MAR-2001; 2001US-0278999P.
PR 27-MAR-2001; 2001US-0279036P.
PR 28-MAR-2001; 2001US-0279344P.
PR 30-MAR-2001; 2001US-0277338P.
PR 30-MAR-2001; 2001US-0279995P.
PR 30-MAR-2001; 2001US-0280233P.
PR 02-APR-2001; 2001US-0280802P.
PR 02-APR-2001; 2001US-0280822P.
PR 02-APR-2001; 2001US-0280900P.
PR 04-APR-2001; 2001US-0281194P.
PR 13-APR-2001; 2001US-0283675P.
PR 30-APR-2001; 2001US-0287424P.
PR 02-MAY-2001; 2001US-0288066P.
PR 03-MAY-2001; 2001US-0288342P.
PR 03-MAY-2001; 2001US-0288528P.
PR 15-MAY-2001; 2001US-0291190P.
PR 16-MAY-2001; 2001US-0291099P.
PR 30-MAY-2001; 2001US-0291240P.
PR 30-MAY-2001; 2001US-0294485P.
PR 31-MAY-2001; 2001US-0294889P.
PR 31-MAY-2001; 2001US-0294999P.
PR 18-JUN-2001; 2001US-0299027P.
PR 19-JUN-2001; 2001US-0299303P.
PR 19-JUN-2001; 2001US-0299310P.
PR 10-JUL-2001; 2001US-0304354P.
PR 31-JUL-2001; 2001US-0309198P.
PR 16-AUG-2001; 2001US-0312903P.
PR 10-SEP-2001; 2001US-0318462P.
PR 12-SEP-2001; 2001US-0318770P.
PR 27-SEP-2001; 2001US-0325430P.
PR 27-SEP-2001; 2001US-0325681P.
PR 18-OCT-2001; 2001US-0330380P.
PR 31-OCT-2001; 2001US-0335301P.
PR 14-NOV-2001; 2001US-0332172P.
PR 14-NOV-2001; 2001US-0332271P.
PR 14-NOV-2001; 2001US-0332272P.
PR 14-NOV-2001; 2001US-0333184P.
PR 14-NOV-2001; 2001US-0333272P.
PR 21-NOV-2001; 2001US-0332094P.
PR 03-DEC-2001; 2001US-0337426P.
PR 03-DEC-2001; 2001US-0338092P.
PR 04-DEC-2001; 2001US-0337185P.
PR 03-JAN-2002; 2002US-0345705P.
PR 07-MAR-2002; 2002US-00092900.
XX
XX (CURA-) CURAGEN CORP.
XX Padigar M, Spytek KA, Shenoy SG, Taupier RJ, Pena CEA, Li L;
XX Zerhusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;
XX Patturajan M, Gangolli E, Vernet CAM, Guo X, Tchernev V;
XX Fernandes ER, Casman SJ, Malyankar UM, Gerlach V, Liu Y, Anderson D;
XX Spaderna SK, Catterton E, Burgess C, Leite M, Zhong H, Alsobrook JP;
XX Lepley DM, Rieger DK;
XX WPI; 2002-723332/78.
XX N-PSDB; ABX97031.
XX
XX NOVX polypeptides and polynucleotides, useful for preventing or treating
XX a disorder associated with aberrant NOVX expression or activity e.g.,
XX cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial

PT asthma.
XX Claim 1; Page 134; 1103pp; English.
XX
CC This invention describes novel human NOVX polypeptides which have
CC cytostatic, cardiant, antiarteriosclerotic, antiasthmatic and hypotensive
CC activity. Pharmaceutical compositions comprising the NOVX proteins or
CC nucleic acid molecules or NOVX antibodies are useful for preventing or
CC treating a disorder associated with aberrant NOVX expression or activity
CC e.g. cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
CC asthma. The products of the invention can be used for gene therapy or in
CC a vaccine. ABU65041-ABU65218 represent the NOVX polypeptides encoded by
CC ABX97008-ABX97185
XX
XX Sequence 616 AA;
SQ
Query Match 45.7%; Score 1412.5; DB 5; Length 616;
Best Local Similarity 46.6%; Pred. No. 7.4e-139;
Matches 292; Conservative 102; Mismatches 171; Indels 61; Gaps 13;
QY 4 LAALAKYWSARRLLVLLVPLALLPILFALPPKEGRCLYVILLMAVYVWCTEALPLSVTA 63
Db 1 MASALSYVSKPKSFVILFVTPLLPLVILMPAKVSCCAYVILMAIYCTEVIPLAVIS 60
QY 64 LLPILFPFMGILPSSKVCQYFLDTNLFSLGLIMASAEERNLHRRALKVLMVGVO 123
Db 61 LMPVLLFPLFQILDQRQVCQYMKDINMLFLGLIVAVAVERNLHKRIALRTLWVGAK 120
QY 124 PARLILGMVTTFSFLSMWLSNTASTAMMLPIASATLK-----SLFGQDTRKDL 172
Db 121 PARMLGFMGVTTALLSMWISNTATTAMVPIVEAILQQMEATSAATEAGLEGQGTINNL 180
QY 173 PREGEDSTAARVGNGLRVTPTMQFLASGEGHAEDVEAPLE-----LPDD 218
Db 181 NLEDDETVKAVLGG--KCVAILSTYVKYKVKQLQINMLTFLKLEQEQDLPGRIPQD 238
QY 219 S---KEEHRNINWGLFSLIPYSASIGGTATLTGAPNLIILGLQKSPFQPC-DVYNFG 274
Db 239 SAQCQEDQERKRLCKAMTLCICVAASIGGTATLTGTGNVVLGQMNELFPDSDKLVNFA 298
QY 275 SWFIFAPFLMLFLVGLMWISFLYGGMSW-RGW-----RKNNSKLQDVAEDKAKAVIOE 328
Db 299 SWFAPFPNNLVMLFPAWLQVLYVNFSSFKKSWGCGLESKKNEK-----AALKVLQE 351
QY 329 EFQNLGPIKFAEQAVFILFCFLFAILLFSRDPKFIQWASL-FAPG---FVSDAVTGVAVI 384
Db 352 EYKLGPLSFAEINVLICFLLVILWFSRDPGMPGMLTVAVWEGETKSYSVATVAIFVA 411
QY 385 TILFFPPSQKPSLKNWDFDKAPNSETE-----PLLSWKKAOETVPWNIIILGGGFA 436
Db 412 TLLFIVPSQPKP---FNFRSQTEGKSPVLIAPPFLDDMKVTKQKVPVGMIVLLGGGFA 467
QY 437 MAKCEESGLSAGWIGQLHLEHVPPLVAVLTIVIAFTEPFASNTATIIIFLPVLAEL 496
Db 468 LAKGSEASGLSVMMGQKQEPHAPPAATLILSLVAVFTECTSNVATITLFLPFASM 527
QY 497 AIRLHVHPLYMLIPGTVSCSYAFMLPVPSTPPNSIAFSTGHLLVKDMVTRTGLLNLKGVLL 556
Db 528 SRSIGLNPLYIMLPCTLSASFAPMLFVATPPNAIVFTYGYHLKVADVMKVTGVIINNIIGVFC 587
QY 557 LSLAMNTWAQAIPLQGTFFPDWAN-TH 581
Db 588 VFLAVNTWGRAIPDLDFPDWANVTH 613
RESULT 7
ABU65063
ID ABU65063 standard; protein; 616 AA.
XX
XX AC ABU65063;
XX
XX XX 20-MAY-2003 (first entry)
XX

Db 468 LAGSEASGLSVWVGKQMEPLHVPAAITLILSLVAVFTECTSNVATTTLFLPIFASM 527
QY 497 AIRLHVHPLYLMPGTVCSCYAFMLPVSTPPNSIAFSTGHLVYKQWVRTGLMNLGVLL 556
Db 528 SRSIGLNPPLYIMLPCTLSASFAMLPVATPPNAIVFTYGHLLKVADWVKGTGVIMNIIGVFC 587
QY 557 LSLAMNTWAQAIIFOLGTFPPDWMAN-TH 581
Db 588 VFLAVNTWGRAIFDLDFHFPDWMANVTH 613

RESULT 8
ADK51048
ID ADK51048 standard; protein; 616 AA.
XX
AC ADK51048;
DT 17-JUN-2004 (first entry)
XX
DE Human NOV18A protein sequence SeqID68.
XX
KW cytostatic; NOVX-agonist; NOVX-antagonist; vaccine; gene therapy; cancer;
KW chromosome mapping; human; NOV18A.
XX
OS Homo sapiens.
XX
PN WO2003083046-A2.
XX
PD 09-OCT-2003.
XX
PF 01-APR-2003; 2003WO-US010142.
XX
PR 02-APR-2002; 2002US-00115479.
PR 05-APR-2002; 2002US-0370349P.
PR 08-APR-2002; 2002US-0370969P.
PR 12-APR-2002; 2002US-0372019P.
PR 22-APR-2002; 2002US-0374379P.
PR 30-MAY-2002; 2002US-0394543P.
PR 03-JUN-2002; 2002US-00160619.
PR 15-AUG-2002; 2002US-0403748P.
PR 04-NOV-2002; 2002US-00287226.
PR 31-MAR-2003; 2003US-00403161.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Anderson DW, Bento P, Boldog FL, Burgess CE, Caeman SJ, Furtak K;
PI Gorman L, Gould-Rothberg BE, Gunther E, Heyes MP, Li L, Spytek KA;
PI Stone DJ, Zhong M, Malyankar UM, Edinger SR, Patturajan M;
PI Rothenberg ME, Smithson G;
XX
DR WPI; 2003-812539/76.
DR N-PSDB; ADK51047.
XX
PT New NOVX polypeptide, useful for preparing a composition for treating or
PT preventing e.g. cancer or for chromosome mapping.
XX
PS Claim 1; SEQ ID NO 68; 433pp; English.
XX
CC This invention relates to novel isolated polypeptides and the DNA
CC sequences which encode them. The invention may be useful for the
CC development of compounds with a cytostatic activity (as NOVX-agonists or
CC antagonists) or vaccines. In addition, the disclosed sequences may be
CC useful for gene therapy. The polypeptide is useful for preparing a
CC composition for treating or preventing a pathological state in a mammal,
CC for example cancer or for chromosome mapping. The present sequence is
CC that of a human NOVX protein of the invention.
XX
SQ Sequence 616 AA;

Query Match 45.7%; Score 1412.5; DB 7; Length 616;
Best Local Similarity 46.6%; Pred. No. 7.4e-139;
Matches 292; Conservative 102; Mismatches 171; Indels 61; Gaps 13;

QY 4 LAALAKYVMSARRLLVLLVPLALPILFALPPKEGRCLYVILLMAVYMWCTEALPLSVTA 63
Db 1 MASALSYVSKPKSFVILFVTPPLLLPLVILMPAKVSCCAYVILLMAVYMWCTEVIPLAVTS 60
QY 64 LPIILPFPFMGILPSPKVCPOYFLDTNLFPLSGLIMASAEIRNLHRRALKVILMLVGVO 123
Db 61 LMPVLLFLPLQLDSRQVCVQYMKDNTNMLFLGGLIVAVAVERNLHRRALKVILMLVGAK 120
QY 124 PARLILGMVVTTSFLSMMLSNASTAMMLPTASATLK-----SLFQORDTRKDL 172
Db 121 PARMLGFMGVTTALLSMMISNTATTAMMVPIVEALQQMEATSAAEAGLEGCGTINNL 180
QY 173 PREGEDSTAARVNGRLRTVPTMOPFLASSEGHAEDVEAPLE-----LPDD 218
Db 181 NAEEDDTVKAVLGG--KCVAILSTVYKVKELQINNLTFLKLEKQEQDGLPGIRPOD 238
QY 219 S---KEEHRHNRNIWGLFISIPYSASIGGTATLTGTAPNLILLGQLKSFPOC-DVWNFG 274
Db 239 SAQCQEDQERKRLCKAMTILCICAAASIGGTATLTGTGNVLLGQMNELFPDSKDLVFA 298
QY 275 SWFIFAFPLMLFLVGLWISFLYGGMSW-RGW-----RKNSKLQDVAEADKAKAVIQE 328
Db 299 SWFAFAPNMLVNLFLFAMLWLFQVYMFSSFKSWGCGLESKNEK-----AALKVLQE 351
QY 329 EFQNLGPIKFAEQAVFLFCLFALLFSRDPKPIPGWASL-FAPG---FVSDAVTGVAVI 384
Db 352 EYRKLGLPLSFAEINVLCFFLLVILWFSRDPGMPGMLTVAVVEGETKSVSDATVAIFVA 411
QY 385 TILFPFSPQKSLKWFDFPKAPNSETE-----PLLSWKKAQETVPNNIILLGGGFA 436
Db 412 TLLFIVPSQPK-----FNFRSQTEGKSPVLIAPPDLDMKVTKQKVPNGI VLLGGGFA 467
QY 437 MAGCEESGLSAWIGGOLHPLHVPPLLAVLITVIAFFTFASNTATIIIFLPVLAE 496
Db 468 LAKGEASGLSVWVGKQMEPLHVPAAITLILSLVAVFTECTSNVATTTLFLPIFASM 527
QY 497 AIRLHVHPLYLMPGTVCSCYAFMLPVSTPPNSIAFSTGHLVYKQWVRTGLMNLGVLL 556
Db 528 SRSIGLNPPLYIMLPCTLSASFAMLPVATPPNAIVFTYGHLLKVADWVKGTGVIMNIIGVFC 587
QY 557 LSLAMNTWAQAIIFOLGTFPPDWMAN-TH 581
Db 588 VFLAVNTWGRAIFDLDFHFPDWMANVTH 613

RESULT 9
ADH42447
ID ADH42447 standard; protein; 616 AA.
XX
AC ADH42447;
XX
DT 25-MAR-2004 (first entry)
XX
DE Novel human protein NOV67d.
XX
KW cardiovascular; antiarteriosclerotic; hypotensive; cytostatic; anorectic;
KW antidiabetic; immunosuppressive; anti-HIV; neuroprotective; nootropic;
KW antiparkinsonian; antiasthmatic; antiinfertility; cardiomyopathy;
KW atherosclerosis; hypertension; cancer; obesity; diabetes; AIDS;
KW multiple sclerosis; graft-versus-host disease; Alzheimer's disease;
KW Parkinson's disease; asthma; fertility disorder; chromosome mapping;
KW tissue typing; preventive medicine; pharmacogenomic; vaccine.
OS Homo sapiens.
XX
PN WO2003102159-A2.
XX
PD 11-DEC-2003.
XX
PF 04-JUN-2003; 2003WO-US017573.
XX
PR 04-JUN-2002; 2002US-0385490P.

XX	04-JUN-2002;	2002US-0385615P.	XX	New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT	04-JUN-2002;	2002US-0385755P.	PT	treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
PT	05-JUN-2002;	2002US-0386041P.	PT	atherosclerosis or diabetes, in chromosome mapping, tissue typing or in
PT	06-JUN-2002;	2002US-0386355P.	PT	pharmacogenomics.
XX	06-JUN-2002;	2002US-0386357P.	XX	Claim 2; SEQ ID NO 1000; 1503pp; English.
PS	06-JUN-2002;	2002US-0386447P.	XX	The invention relates to 566 new isolated human polypeptides and their
CC	06-JUN-2002;	2002US-0386459P.	CC	encoding genes, sequences that are at least 95% identical to these or
CC	06-JUN-2002;	2002US-0386465P.	CC	sequences comprising one or more conservative substitutions in these. The
CC	06-JUN-2002;	2002US-0386684P.	CC	polypeptide, polynucleotide and antibodies against the polypeptides are
CC	07-JUN-2002;	2002US-0386701P.	CC	useful in diagnosing, treating or preventing NOVX-associated disorders,
CC	07-JUN-2002;	2002US-0386796P.	CC	e.g. cardiomyopathy, atherosclerosis, hypertension, cancer, obesity,
CC	07-JUN-2002;	2002US-0386931P.	CC	diabetes, AIDS, multiple sclerosis, graft-versus-host disease,
CC	07-JUN-2002;	2002US-0387081P.	CC	Alzheimer's disease, Parkinson's disease, asthma, or fertility disorders.
CC	07-JUN-2002;	2002US-0387083P.	CC	The nucleic acids are further used as hybridization probes, in chromosome
CC	10-JUN-2002;	2002US-0387429P.	CC	mapping, tissue typing, preventive medicine, and pharmacogenomics. The
CC	10-JUN-2002;	2002US-0387540P.	CC	polypeptides are also useful as vaccines. This sequence represents an
CC	10-JUN-2002;	2002US-0387866P.	CC	example of the polypeptide of the invention.
CC	10-JUN-2002;	2002US-0387868P.	XX	Sequence 616 AA;
CC	11-JUN-2002;	2002US-0387606P.	XX	Query Match 45.7%; Score 1412.5; DB 8; Length 616;
CC	11-JUN-2002;	2002US-0387610P.	Best Local Similarity 46.6%; Pred. No. 7.4e-139;	
CC	11-JUN-2002;	2002US-0387659P.	Matches 292; Conservative 102; Mismatches 171; Indels 61; Gaps 13;	
CC	11-JUN-2002;	2002US-0387668P.		
CC	11-JUN-2002;	2002US-0387696P.		
CC	11-JUN-2002;	2002US-0387859P.		
CC	12-JUN-2002;	2002US-0387866P.		
CC	12-JUN-2002;	2002US-0387868P.		
CC	12-JUN-2002;	2002US-0387934P.		
CC	12-JUN-2002;	2002US-0387960P.		
CC	12-JUN-2002;	2002US-0388022P.		
CC	12-JUN-2002;	2002US-0388066P.		
CC	12-JUN-2002;	2002US-0388432P.		
CC	12-JUN-2002;	2002US-0388479P.		
CC	13-JUN-2002;	2002US-0389123P.		
CC	14-JUN-2002;	2002US-0389120P.		
CC	14-JUN-2002;	2002US-0389146P.		
CC	17-JUN-2002;	2002US-0389742P.		
CC	18-JUN-2002;	2002US-0389604P.		
CC	18-JUN-2002;	2002US-0389884P.		
CC	19-JUN-2002;	2002US-0390066P.		
CC	19-JUN-2002;	2002US-0390144P.		
CC	19-JUN-2002;	2002US-0390209P.		
CC	25-JUN-2002;	2002US-0391726P.		
CC	06-AUG-2002;	2002US-0401628P.		
CC	09-AUG-2002;	2002US-0402268P.		
CC	12-AUG-2002;	2002US-0402822P.		
CC	13-AUG-2002;	2002US-0403458P.		
CC	15-AUG-2002;	2002US-0403617P.		
CC	15-AUG-2002;	2002US-0403732P.		
CC	26-AUG-2002;	2002US-0406182P.		
CC	12-SEP-2002;	2002US-0410083P.		
CC	13-SEP-2002;	2002US-0410505P.		
CC	23-SEP-2002;	2002US-0412955P.		
CC	30-SEP-2002;	2002US-0415195P.		
CC	23-OCT-2002;	2002US-0420627P.		
CC	23-OCT-2002;	2002US-0420718P.		
CC	24-OCT-2002;	2002US-0420852P.		
CC	31-OCT-2002;	2002US-0422750P.		
CC	01-NOV-2002;	2002US-0423095P.		
CC	05-NOV-2002;	2002US-0423748P.		
XX	(CURA-) CURAGEN CORP.			
PA				
XX	Alsobrook JP, Anderson DW, Baumgartner JC, Berghs C, Boldog FL;			
PI	Burgess CE, Casman SJ, Catterton E, Dhanabal M, Edinger SR;			
PI	Ellerman K, Ettenberg S, Gangolli EA, Gerlach VL, Gorman L;			
PI	Groese WM, Gunther E, Guo X, Gusev VY, Herrmann JL, Ji W, Kekuda R;			
PI	Khrantsov NV, Larochelel WJ, Li L, Liang H, Low K, Macdougall JR;			
PI	MacLachlan T, Malayanar K, McQueeney K, Mezick AJ, Miller CE;			
PI	Millet I, Padigar M, Patturajan M, Peyman JA, Qian X, Raetelli L;			
PI	Rieger DK, Rothenberg ME, Shenoy SG, Shinkets RA, Smithson G;			
PI	Spytek KA, Stone DJ, Sukumaran S, Szekeres ES, Vernet CAM, Voss EZ;			
PI	Wolenc AR, Zhong M, Zhong H;			
XX	WPI; 2004-053467/05.			
DR	N-PSDB; ADH42446.			

ADN61777
ID ADN61777 standard; protein; 616 AA.
XX
AC ADN61777;
XX
DT 01-JUL-2004 (first entry)
XX
DE Human novel protein NOV14b.
XX
KW Human; NOVX; diabetes; obesity; infectious disease; anorexia;
KW cancer-associated cachexia; cancer; neurodegenerative disorder;
KW Alzheimer's disease; Parkinson's disease; immune disorder;
KW haematopoietic disorder; dyslipidaemia; chronic disease.
XX
OS Homo sapiens.
XX
PN US2004043382-A1.
XX
PD 04-MAR-2004.
XX
PF 07-MAR-2002; 2002US-00092900.
XX
PR 08-MAR-2001; 2001US-0274191P.
PR 08-MAR-2001; 2001US-0274194P.
PR 08-MAR-2001; 2001US-0274281P.
PR 08-MAR-2001; 2001US-0274322P.
PR 09-MAR-2001; 2001US-0274849P.
PR 12-MAR-2001; 2001US-0275235P.
PR 13-MAR-2001; 2001US-0275578P.
PR 13-MAR-2001; 2001US-0275579P.
PR 13-MAR-2001; 2001US-0275601P.
PR 14-MAR-2001; 2001US-0276000P.
PR 16-MAR-2001; 2001US-0276776P.
PR 19-MAR-2001; 2001US-0276994P.
PR 20-MAR-2001; 2001US-0277239P.
PR 20-MAR-2001; 2001US-0277321P.
PR 20-MAR-2001; 2001US-0277327P.
PR 20-MAR-2001; 2001US-0277338P.
PR 21-MAR-2001; 2001US-0277791P.
PR 22-MAR-2001; 2001US-0277833P.
PR 23-MAR-2001; 2001US-0278152P.
PR 26-MAR-2001; 2001US-0278894P.
PR 27-MAR-2001; 2001US-0278999P.
PR 27-MAR-2001; 2001US-0279036P.
PR 28-MAR-2001; 2001US-0279344P.
PR 30-MAR-2001; 2001US-0279995P.
PR 30-MAR-2001; 2001US-0280233P.
PR 02-APR-2001; 2001US-0280802P.
PR 02-APR-2001; 2001US-0280822P.
PR 02-APR-2001; 2001US-0280900P.
PR 04-APR-2001; 2001US-0281444P.
PR 13-APR-2001; 2001US-0283675P.
PR 30-APR-2001; 2001US-0287424P.
PR 02-MAY-2001; 2001US-0288066P.
PR 03-MAY-2001; 2001US-0288342P.
PR 03-MAY-2001; 2001US-0288528P.
PR 15-MAY-2001; 2001US-0291130P.
PR 16-MAY-2001; 2001US-0291099P.
PR 30-MAY-2001; 2001US-0291240P.
PR 31-MAY-2001; 2001US-0294485P.
PR 31-MAY-2001; 2001US-0294889P.
PR 31-MAY-2001; 2001US-0294899P.
PR 18-JUN-2001; 2001US-0299027P.
PR 19-JUN-2001; 2001US-0299303P.
PR 19-JUN-2001; 2001US-0299310P.
PR 10-JUL-2001; 2001US-0304354P.
PR 31-JUL-2001; 2001US-0309198P.
PR 16-AUG-2001; 2001US-0312903P.
PR 10-SEP-2001; 2001US-0318462P.
PR 12-SEP-2001; 2001US-0318770P.
PR 27-SEP-2001; 2001US-0325430P.
PR 27-SEP-2001; 2001US-0325681P.
PR 18-OCT-2001; 2001US-0330380P.

PR 31-OCT-2001; 2001US-0335301P.
PR 14-NOV-2001; 2001US-0332172P.
PR 14-NOV-2001; 2001US-0332271P.
PR 14-NOV-2001; 2001US-0332272P.
PR 14-NOV-2001; 2001US-0333184P.
PR 14-NOV-2001; 2001US-0333272P.
PR 21-NOV-2001; 2001US-0332094P.
PR 03-DEC-2001; 2001US-0337426P.
PR 03-DEC-2001; 2001US-0338092P.
PR 04-DEC-2001; 2001US-0337185P.
PR 03-JAN-2002; 2002US-0345705P.
XX
PA (PADI/) PADIGARU M.
PA (SPYT/) SPYTEK K A.
PA (SHEN/) SHENOY S G.
PA (TAUP/) TAUPIER R J.
PA (PENA/) PENNA C E A.
PA (LILL/) LI L.
PA (ZERH/) ZERHUSEN B D.
PA (GUSE/) GUSEV V Y.
PA (JIWV/) JI W.
PA (GORM/) GORMAN L.
PA (MILL/) MILLER C E.
PA (KEKU/) KEKUDA R.
PA (PATT/) PATTURAJAN M.
PA (GANG/) GANGOLLI E A.
PA (VERN/) VERNET C A M.
PA (GUOX/) GUO X S.
PA (TCHE/) TCHERNEV V T.
PA (FERN/) FERNANDES E R.
PA (CASM/) CASMAN S J.
PA (MALX/) MALYANKAR U M.
PA (GERL/) GERLACH V.
PA (LIUY/) LIU Y.
PA (ANDE/) ANDERSON D W.
PA (SPAD/) SPADERNA S K.
PA (CATT/) CATTERTON E.
PA (LEIT/) LEITE M W.
PA (ZHON/) ZHONG H.
PA (ALSO/) ALSOBROOK J P.
PA (LEPL/) LEPLEY D M.
PA (RIEG/) RIEGER D K.
PA (BURG/) BURGESS C E.
XX
PI Padigaru M, Spytek KA, Shenoy SG, Taupier RJ, Pena CEA, Li L;
PI Zehrusen BD, Gusev VY, Ji W, Gorman L, Miller CE, Kekuda R;
PI Patturajan M, Gangolli EA, Vernet CAM, Guo XS, Tchernev VT;
PI Fernandes ER, Casman SJ, Malyankar UM, Gerlach V, Liu Y;
PI Anderson DW, Spaderna SK, Catterton E, Leite MW, Zhong H;
PI Alsobrook JP, Lepley DM, Rieger DK, Burgess CE;
XX
WPI; 2004-225693/21.
DR N-PSDB; ADN61776.
XX
PT New NOVX polypeptides and nucleic acid molecules useful for diagnosing,
PT preventing or treating NOVX-associated disorders, e.g. cancer, diabetes,
PT infection or obesity, and in chromosome mapping, tissue typing or
PT pharmacogenomics.
XX
PS Claim 1; SEQ ID NO 46; 786pp; English.
XX
CC The invention relates to an isolated polypeptide (designated NOVX, or
CC NOV1-NOV127) comprising a sequence selected from 178 fully defined amino
CC acid sequences (and their mature forms, variants and fragments). Also
CC included are an isolated nucleic acid molecule encoding NOVX, a vector
CC comprising the nucleic acid, a cell comprising the vector, methods for
CC determining the presence or amount of the polypeptide or the nucleic acid
CC molecule in a sample, methods for determining the presence of or
CC predisposition to a disease associated with altered levels of expression
CC of the above polypeptide or nucleic acid molecule in a first mammalian
CC subject, a method for identifying an agent that binds to the above
CC polypeptide, a method for identifying a potential therapeutic agent for
CC use in the treatment of a pathology that is related to aberrant

PR	03-DEC-2001; 2001US-0338092P.	CC	dyslipidaemias, and other chronic diseases. These may also be used in
PR	04-DEC-2001; 2001US-0337185P.	CC	chromosome mapping, tissue typing, preventive medicine and
PR	03-JAN-2002; 2002US-0345705P.	CC	pharmacogenomics. The polypeptides are also useful as vaccines. The
XX		CC	present sequence represents a NOVX protein of the invention.
XX		XX	Sequence 616 AA;
PA	(PADI/) PADIGARU M.	Query Match	45.7%; Score 1412.5; DB 8; Length 616;
PA	(SPYT/) SPYTEK K A.	Best Local Similarity	46.6%; Pred. No. 7.4e-139;
PA	(SHEN/) SHENOY S G.	Matches	292; Conservative 102; Mismatches 171; Indels 61; Gaps 13;
PA	(TAUP/) TAUPIER R J.		
PA	(PENA/) PENA C E A.		
PA	(LILL/) LI L.		
PA	(ZERH/) ZERHUSEN B D.		
PA	(GUSE/) GUSEV V Y.		
PA	(JIMW/) JI W.		
PA	(GORM/) GORMAN L.		
PA	(MILL/) MILLER C E.		
PA	(KEGU/) KEKUDA R.		
PA	(PATT/) PATTURAJAN M.		
PA	(GANG/) GANGOLLI E A.		
PA	(VERN/) VERNET C A M.		
PA	(GUOX/) GUO X S.		
PA	(TCHE/) TCHERNEV V T.		
PA	(FERN/) FERNANDES E R.		
PA	(CASM/) CASMAN S J.		
PA	(MALY/) MALYANKAR U M.		
PA	(GERL/) GERLACH V.		
PA	(LIUY/) LIU Y.		
PA	(ANDE/) ANDERSON D W.		
PA	(SPAD/) SPADERNA S K.		
PA	(CATT/) CATTERTON E.		
PA	(LEIT/) LEITE M W.		
PA	(ZHON/) ZHONG H.		
PA	(ALSO/) ALSOBROOK J P.		
PA	(LEPL/) LEPLLEY D M.		
PA	(RIEG/) RIEGER D K.		
PA	(BURG/) BURGESS C E.		
XX			
PI	Padigaru M, Spyttek KA, Shenoy SG, Taupier RJ, Pena CE, Li L;		
PI	Zerhusen BD, Gusev VV, Ji W, Gorman L, Miller CE, Kekuda R;		
PI	Patturajan M, Gangolli EA, Vernet CM, Guo XS, Tchernev VT;		
PI	Fernandes ER, Casman SJ, Malyankar UM, Gerlach V, Liu Y;		
PI	Anderson DW, Spaderna SK, Catterton E, Leite MW, Zhong H;		
PI	Alsobrook JP, Lepley DM, Rieger DK, Burgess CE;		
XX			
DR	WPI: 2004-225693/21.		
DR	N-ESDB; ADN61778.		
XX			
PT	New NOVX polypeptides and nucleic acid molecules useful for diagnosing,		
PT	preventing or treating NOVX-associated disorders, e.g. cancer, diabetes,		
PT	infection or obesity, and in chromosome mapping, tissue typing or		
PT	pharmacogenomics.		
XX			
PS	Claim 1; SEQ ID NO 48; 786pp; English.		
XX			
CC	The invention relates to an isolated polypeptide (designated NOVX, or		
CC	NOV1-NOV127) comprising a sequence selected from 178 fully defined amino		
CC	acid sequences (and their mature forms, variants and fragments). Also		
CC	included are an isolated nucleic acid molecule encoding NOVX, a vector		
CC	comprising the nucleic acid, a cell comprising the vector, methods for		
CC	determining the presence or amount of the polypeptide or the nucleic acid		
CC	molecule in a sample, methods for determining the presence of or		
CC	predisposition to a disease associated with altered levels of expression		
CC	of the above polypeptide or nucleic acid molecule in a first mammalian		
CC	subject, a method for identifying an agent that binds to the above		
CC	polypeptide, a method for identifying a potential therapeutic agent for		
CC	use in the treatment of a pathology that is related to aberrant		
CC	expression or physiological interactions of the polypeptide, a method of		
CC	screening for a modulator of activity or of latency or predisposition to		
CC	a pathology associated with the polypeptide and a method for modulating		
CC	the activity of the polypeptide cited above. The composition and methods		
CC	are useful for diagnosing, preventing or treating diseases such as		
CC	diabetes, obesity, infectious diseases, anorexia, cancer-associated		
CC	cachexia, cancer, neurodegenerative disorders like Alzheimer's disease or		
CC	Parkinson's disease, immune disorders, haematopoietic disorders,		

FT Domain /note= "Casein kinase II phosphorylation site"
406..422
FT Domain /note= "Transmembrane domain"
409..514
FT Domain /note= "Transmembrane domain"
423..428
FT Modified-site
440..459
FT Domain /note= "N-myristoylation site"
466..483
FT Domain /note= "Transmembrane domain"
498..514
FT Domain /note= "Transmembrane domain"
528..544
FT Modified-site
562..565
FT Modified-site /note= "N-glycosylation site"
564..567
FT /note= "Casein kinase II phosphorylation site"
US2002193582-A1.

19-DEC-2002.
17-JUN-2002; 2002US-00173519.
18-JUN-2001; 2001US-0298970P.
(MILL-) MILLENNIUM PHARM INC.
Curtis RAJ;
WPI; 2003-644620/61.
N-PSDB; AAD58822.

Novel 69624 polypeptide, a human transporter family member, useful for treating disorders e.g. hypocitranuria, such as hypocitranuria, formation of calcium stones, mental retardation abnormal body sulfate homeostasis.
Claim 8; Page 43-45; Opp; English.

The invention relates to 69624 polypeptide, a human transporter family member and its corresponding nucleic acid. 69624 protein is useful for developing novel diagnostic and therapeutic agents for 69624-mediated or related disorders. 69624 protein act as therapeutic or diagnostic agents for renal, neurological, colon or hepatic disorders. It act as diagnostic targets and therapeutic agents for treating disorders such as hypocitranuria, formation of calcium stones, mental retardation (Canavan disease) or abnormal body sulphate homeostasis. 69624 protein may act as diagnostic targets and therapeutic agents for controlling cellular proliferative and/or differentiative disorders such as carcinoma, sarcoma, metastatic disorder or haematopoietic neoplastic disorders e.g., leukaemia, immune disorders such as autoimmune disorders (diabetes mellitus, arthritis), multiple sclerosis, encephalomyelitis, myasthenia gravis, systemic lupus erythematosus, cardiovascular disorders such as arteriosclerosis, atherosclerosis, ischaemia reperfusion injury, cardiac hypertrophy, atrial fibrillation etc; and disorders involving abnormal or excessive pain. 69624 sequence is useful as pharmacodynamic marker and is also used in forensic identification of a biological sample. The present sequence is human 69624 protein

Sequence 568 AA;

Query Match 45.6%; Score 1411.5; DB 7; Length 568;
Best Local Similarity 47.1%; Pred. No. 8.3e-139;
Matches 286; Conservative 100; Mismatches 150; Indels 71; Gaps 13;

QY 4 LAALAKVWSARRLLVLLPLALLPLFALPKGKGLVYLIMAYVWCTEALPLSVTA 63
Db 1 MASALSVSFKSVILFVPLLLPLVILPAKVFRCVAILMAYVWCTEIVPLAVTS 60
QY 64 LLPILFPFMGILPSSKVCQYFDLTNPLFLSLGILMASAIEERNLHRIALKVLMVGQV 123

Db 61 LMPVLLFPLFQILDSRQVCVQYMKDNTNMLFLGGLIVAVAVERNLHKRIALRTLLVWGAK 120
QY 124 PARLLILGMVTTTSLFSLMWSLNTASTAMMLPIASAILKSLFGQDTRKDLPRGEDSTA 183
Db 121 PARMLILGMVGTALLSLMWSLNTASTAMMPIVEAILQQM-----EATSAA 165
QY 184 RGNGLRTVPTMQFLASSEGHAEDVEAPLELPDD-----SKEBEHRNRIWKGLF 233
Db 166 TEAGLELV-----DKGKAK-----ELPGSQVIEGPTLGOQEQERKRLCKAMT 209
QY 234 ISIPYSASIGGTATLTGTAPNLIILGQLKSPFPQC-DVNVFGSWFIFAPPLMLLFLVGV 292
Db 210 LCICYAASIGGTATLTGTGNVLLQNMNELFPDSKDLVNFASFAPFAPNMLVMLLFAW 269
QY 293 LMSISLYGCMWSW-RGW-----RKQSKLODVAEDKAKAVIOBEFQNLGPIKPAEQAVFIL 346
Db 270 LMLQFVYMRFPFKKSGCGLSKNEK-----AALKVLEBYRKLGLPLSFAEINVLIC 332
QY 347 FCILFALLFSRDPKFTIPGWASL-FAPG----FVSDAVTGAIVTILFFFPKPSLKWFD 402
Db 323 FFLVILWFSDRPGFMPGMLTVANVEGETKYVSDATVAIFVATLLFIVFSQKPK----FN 378
QY 403 FKAPNSET-----PLLSWKKAQETVPWNILLLGGGFAMAKGCEESGLSWIGQLH 455
Db 379 FRSQTEERKTPFPYPPPLLDWKVTQEKVPWIGVLLGGGFALAKGSEASGLSYVMGKQME 438
QY 456 ELEHVPPLLAVALITVYIAFFTEFASNTATIIIFLPVLAELAIRLHVHPLVLMIPGTVC 515
Db 439 PLHAVPPAAITILSLVAVFTECTSNVATITLFLPFIASMSRSGISGLNPLIIMPLCTLSA 498
QY 516 SYAFMLPVSPTPPNSIAFSTGHLVVKDMVRTGLIMNLMGLVLLSLAMNTWAQAI 575
Db 499 SFAPMLPVATPPNAIVFTYGHKLVADVMKVTGYIMNIIGVFCVFLAVNTWGRAIFDL 558
QY 576 DWAN-TH 581
Db 559 DWANVTH 565
RESULT 14
ADKS1052
ID ADKS1052 standard; protein; 568 AA.
AC ADKS1052;
XX
DT 17-JUN-2004 (first entry)
XX
DE Human NOV18C protein sequence SeqID72.
XX
KW Cytostatic; NOVX-agonist; NOVX-antagonist; vaccine; gene therapy; cancer;
KW chromosome mapping; human; NOV18C.
XX
OS Homo sapiens.
XX
PN WO2003083046-A2.
PD
XX 09-OCT-2003.
PF
XX 01-APR-2003; 2003WO-US010142.
XX
PR 02-APR-2002; 2002US-00115479.
PR 05-APR-2002; 2002US-0370349P.
PR 08-APR-2002; 2002US-0370969P.
PR 12-APR-2002; 2002US-0372019P.
PR 22-APR-2002; 2002US-0374379P.
PR 30-MAY-2002; 2002US-0384543P.
PR 03-JUN-2002; 2002US-00160619.
PR 15-AUG-2002; 2002US-0403748P.
PR 04-NOV-2002; 2002US-00287226.
PR 31-MAR-2003; 2003US-00403161.
XX (CURA-) CURAGEN CORP.
XX

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2005, 08:36:57 ; Search time 23.069 Seconds
(without alignments)
1941.542 Million cell updates/sec

Title: US-10-017-479A-5
Perfect score: 3094
Sequence: 1 MAALAAAKVWSARLLVL.....HAANVTALPPALTNTVTQL 600

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : Issued Patents AA:
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTCUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1333.5	43.1	599	4	US-09-949-016-9866
2	1208.5	39.1	601	4	US-09-949-016-9977
3	1188	38.4	627	4	US-09-556-916-2
4	1186	38.3	627	4	US-09-949-016-6840
5	1183.5	38.3	626	4	US-09-556-916-14
6	1182	38.2	627	4	US-09-556-916-4
7	1180	38.1	627	4	US-09-556-916-8
8	1177.5	38.1	626	4	US-09-556-916-16
9	1175.5	38.0	626	4	US-09-556-916-20
10	1174	37.9	627	4	US-09-556-916-10
11	1169.5	37.8	626	4	US-09-556-916-22
12	1019	32.9	581	4	US-09-556-916-6
13	1014.5	32.8	580	4	US-09-556-916-18
14	1011	32.7	581	4	US-09-556-916-12
15	1006.5	32.5	580	4	US-09-556-916-24
16	950	30.7	561	4	US-09-949-016-8161
17	813	26.3	233	4	US-09-489-847-176
18	740	23.9	527	4	US-09-602-787A-516
19	517	16.7	368	4	US-09-270-767-42029
20	501	16.2	470	4	US-09-543-681A-5952
21	461.5	14.9	335	4	US-09-602-787A-518
22	392.5	12.7	524	3	US-09-134-001C-5457
23	308.5	10.0	180	4	US-09-270-767-42669
24	307.5	9.9	421	4	US-09-248-796A-20749
25	265.5	8.6	230	4	US-09-270-767-43713
26	255.5	8.3	194	4	US-09-270-767-59102
27	251	8.1	132	4	US-09-270-767-57987

28	234	7.6	547	4	US-09-489-039A-13843	Sequence 13843, A
29	207.5	6.7	624	4	US-09-543-681A-4343	Sequence 4343, Ap
30	206	6.7	169	4	US-09-270-767-57286	Sequence 57286, A
31	204	6.6	494	4	US-09-543-681A-7033	Sequence 7033, Ap
32	203	6.6	302	4	US-09-902-540-10445	Sequence 10445, A
33	189.5	6.1	614	4	US-09-489-039A-12605	Sequence 12605, A
34	189	6.1	596	4	US-09-902-540-13547	Sequence 13547, A
35	185	6.0	548	4	US-09-902-540-11870	Sequence 11870, A
36	167	5.4	694	4	US-09-252-991A-22637	Sequence 22637, A
37	164.5	5.3	470	4	US-09-438-185A-209	Sequence 209, App
38	161.5	5.2	493	4	US-09-540-236-2120	Sequence 2120, Ap
39	157	5.1	443	4	US-09-602-787A-532	Sequence 532, App
40	142.5	4.6	430	3	US-09-134-001C-2981	Sequence 2981, Ap
41	141	4.6	478	3	US-09-134-001C-4637	Sequence 4637, Ap
42	137	4.4	441	4	US-09-489-039A-10612	Sequence 10612, A
43	130.5	4.2	597	4	US-09-489-039A-14112	Sequence 14112, A
44	122	3.9	167	4	US-09-248-796A-20740	Sequence 20740, A
45	120.5	3.9	363	4	US-09-270-767-42613	Sequence 42613, A
46	115.5	3.7	319	4	US-09-270-767-57927	Sequence 57927, A
47	114	3.7	429	4	US-08-311-731A-287	Sequence 287, App
48	113	3.7	437	4	US-09-543-681A-6984	Sequence 6984, Ap
49	113	3.7	1042	4	US-09-252-991A-30444	Sequence 30444, A
50	112.5	3.6	1077	4	US-09-412-210-1	Sequence 1, Appli
51	112.5	3.6	1077	4	US-10-121-911A-1	Sequence 1, Appli
52	111	3.6	211	4	US-09-198-452A-1166	Sequence 1166, Ap
53	111	3.6	381	4	US-09-710-279-3284	Sequence 3284, Ap
54	110.5	3.6	3421	4	US-09-452-638-53	Sequence 53, Appli
55	110	3.6	494	4	US-09-603-208A-232	Sequence 232, App
56	108.5	3.5	449	4	US-09-328-352-7512	Sequence 7512, Ap
57	108.5	3.5	554	4	US-09-543-681A-5774	Sequence 5774, Ap
58	108	3.5	509	4	US-09-134-000C-5949	Sequence 5949, Ap
59	108	3.5	611	2	US-08-677-049-2	Sequence 2, Appli
60	106.5	3.4	580	4	US-09-657-252-2	Sequence 2, Appli
61	106	3.4	540	3	US-09-433-994-2	Sequence 2, Appli
62	105	3.4	574	4	US-09-902-540-10370	Sequence 10370, A
63	104.5	3.4	456	3	US-09-058-389A-2	Sequence 2, Appli
64	104.5	3.4	456	3	US-09-611-781-2	Sequence 2, Appli
65	104.5	3.4	457	4	US-09-949-016-7211	Sequence 7211, Ap
66	104.5	3.4	482	4	US-09-328-352-7784	Sequence 7784, Ap
67	104.5	3.4	1912	4	US-09-495-714C-2	Sequence 2, Appli
68	104.5	3.4	1977	4	US-09-495-714C-4	Sequence 4, Appli
69	104	3.4	502	4	US-09-252-991A-23327	Sequence 23327, A
70	104	3.4	535	4	US-09-107-433-3342	Sequence 3342, Ap
71	104	3.4	540	4	US-09-583-110-3359	Sequence 3359, Ap
72	104	3.4	549	4	US-09-489-039A-13779	Sequence 13779, A
73	103	3.3	116	4	US-09-602-787A-520	Sequence 520, App
74	103	3.3	345	4	US-09-902-540-9836	Sequence 9836, Ap
75	103	3.3	514	4	US-09-543-681A-5352	Sequence 5352, Ap
76	103	3.3	594	4	US-09-252-991A-27335	Sequence 27335, A
77	102.5	3.3	598	4	US-09-252-991A-28599	Sequence 28599, A
78	102	3.3	505	4	US-09-328-352-7470	Sequence 7470, Ap
79	102	3.3	1014	4	US-09-252-991A-29868	Sequence 29868, A
80	101	3.3	395	4	US-09-543-681A-6203	Sequence 6203, Ap
81	100.5	3.2	562	4	US-09-489-039A-8574	Sequence 8574, Ap
82	100	3.2	344	4	US-08-311-731A-172	Sequence 172, App
83	100	3.2	518	4	US-09-252-991A-30478	Sequence 30478, A
84	100	3.2	532	4	US-09-252-991A-22421	Sequence 22421, A
85	100	3.2	693	4	US-09-949-016-9666	Sequence 9666, Ap
86	99.5	3.2	279	4	US-09-602-787A-658	Sequence 658, App
87	99.5	3.2	326	3	US-09-058-389A-3	Sequence 3, Appli
88	99.5	3.2	326	3	US-09-611-781-3	Sequence 3, Appli
89	99.5	3.2	408	3	US-09-134-001C-3410	Sequence 3410, Ap
90	99.5	3.2	463	4	US-09-583-110-5223	Sequence 5223, Ap
91	99.5	3.2	878	4	US-09-438-185A-729	Sequence 729, App
92	99	3.2	534	4	US-09-252-991A-30725	Sequence 30725, A
93	98.5	3.2	300	4	US-09-107-532A-5624	Sequence 5624, Ap
94	98.5	3.2	313	4	US-09-902-540-15660	Sequence 15660, A
95	98.5	3.2	378	4	US-09-055-097-1	Sequence 1, Appli
96	98.5	3.2	378	4	US-09-373-902-1	Sequence 1, Appli
97	98.5	3.2	393	4	US-09-949-016-11567	Sequence 11567, A
98	98	3.2	525	4	US-09-902-540-15103	Sequence 15103, A
99	98	3.2	767	4	US-09-540-236-2346	Sequence 2346, Ap
100	97.5	3.2	378	4	US-09-831-630-13	Sequence 13, Appli

ALIGNMENTS

```

RESULT 1
US-09-949-016-9866
; Sequence 9866, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9866
; LENGTH: 599
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-9866

Query Match      43.1%; Score 1333.5; DB 4; Length 599;
Best Local Similarity 45.4%; Pred. No. 9.1e-126;
Matches 280; Conservative 109; Mismatches 175; Indels 53; Gaps 13;

Qy 4 LAALAKKWSARRLLVLLPLALLPILFALPKPEGRCLYVILLMAYVWCTEALPLSVTA 63
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 8 MATCQALWAVRSYLVFFVPIILLPLILVPSKEAYCAVAILMALFWCTEALPLAVTA 67
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 64 LPLILFPFGIILPSSKVCQYFDLTNPLFLGLINASAEERNLHRRALKVLMVGQV 123
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 68 LPLILFPFGIILPSSKVCQYFDLTNPLFLGLINASAEERNLHRRALKVLMVGQV 127
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 124 PARLILGMVTTSLMNLNTASTAMMLPIASAILKSLFG-----QDTRKDLPRGEDSTAAV 183
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 128 PAPILGLFVLTAFLSMWISNTATSMVPIAHVLDQ-----HSSQASSNVE 176
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 184 RGNGLRTV-----PTFMQFLASSEGHAEDVEAPLELPPDDKEEHRHNIWKGLF----- 233
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 177 EGSNNPTFELQEPSPQKEVTKLDNGQA-----LPVTSASSEGRAHLSQKHLHLTQCM 228
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 234 -ISIPYSASIGGTATLTGTAPNLILGLQKSPFPQCDVNVFGSWFIAPFLMLLLVGV 291
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 229 SLVCYASIGGIATLTGTAPNLVLOGINSLFPQNGNVNFAVSWFAPPTWVILLLLA 288
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 292 WLWISFLYGGMSWRGKKN-----SKLDVAEDKAKAVIQEENLGPRIKFAEQAVILF 347
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 289 WLWQIILFNGFR-----KNFGICEKQEQ-QQQAAYCVIQTREHLLGPMFAEKAISILF 343
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 348 CLFAILLFSRDPKPIPGWASLFAF-----GPFVDAVTGVAIVTILFPFPKSLKMWFD 402
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 344 VILVLLWFTREPGFPLGNGNLAFNAKGSMSVSGTVAIFGIIMFIPSKFGLTQ--D 401
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 403 FKPNSETEP-----LLSWKKAQETVPMNIIILGGGFAMAKGCEESGLSAWIGQLHLEHV 460
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 402 PENFGKLPAGLGLDWTQNMKNPWNIVLLGGGYALAKGSEGLSEWLNKUTPLQSV 461
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 461 PPLAVLLITVIAFFTEFASNTATIIIFLPVLAELAIRLHVHPLYLMIPTVSCSYAFM 520
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 462 PAPAIAILSLVATFECTSNVATITIFLPILASMAQAICLHPLYVWMLPCTLATSLAFM 521
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 521 LPVSTPPNSIAFSTGHLVKKDMVTGLLMNMVGLLSLMMNTWAQAIIFQLGTFPDWANT 580
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 522 LPVATPPNAIVFSFGDLKVLDMARAGFLNIIIGVLIILALAINSWGIFLPSLHSPSQAQ- 580
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

Qy 581 HAANVTA-LPPALTNNNT 596
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 581 --SNTTAQCLPSLANTT 595
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 2
US-09-949-016-9977
; Sequence 9977, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9977
; LENGTH: 601
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-9977

Query Match      39.1%; Score 1208.5; DB 4; Length 601;
Best Local Similarity 40.8%; Pred. No. 4e-113;
Matches 250; Conservative 119; Mismatches 186; Indels 57; Gaps 12;

Qy 15 RRLVLLVPLALLPILFALPKPEGRCLYVILLMAYVWCTEALPLSVTALLPILPFG 74
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 17 RRFLFVFTVLLPLPLVILTKAECAVTLFVATFWLTALPLSVTALLPILPFG 76
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 75 ILPSSKVCQYFDLTNPLFLGLINASAEERNLHRRALKVLMVGQVAPARLILGMV 134
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 77 IMPSKVASAYPKDFHLLIGVICLATSIEKNLHRRALKVLMVGQVAPARLILGMV 136
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 135 TFLSMLNLTASTAMMLPIASAILKSLFG-----QDTRKDLPRGEDSTAAV 186
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 137 TFLSMLNLTASTAMMLPIASAILKSLFG-----QDTRKDLPRGEDSTAAV 196
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 187 GL-----RTVPTMQFLASSEGHAED-----VEAPLELPPDDKEEHRHNIWKGLF----- 233
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 197 EIMERKEKTPV-----GYNNDTGKISKVELEKNSGMRTKYRT-KKHVTRKLT 246
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 234 -ISIPYSASIGGTATLTGTAPNLILGLQKSPFPQCDVNVFGSWFIAPFLMLLLVGV 292
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 247 CLCIAVSSITIGLTTITGTSTNLIPAEYFNTRYPCDCLNFGSWFTSFPAALILL 306
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 293 LMTISFLYGGMSWRGKKNKQDVAEDKAKAVIQEENLGPRIKFAEQAVILFCLFAI 352
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 307 IWLQWFLFNGFR-MFKCGKTKVQKCAEVIKQEQKGLPIRYQEIIVTLVLFIMAL 365
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 353 LLSRDPKPIPGWASLFAF-----GPFVDAVTGVAIVTILFPFPKSLKMWFD 410
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 366 LWSRDPGPGVPGNSALFSEYPGPATDS-VALLIGLFLFLIPAKTLTK-----TPTGEI 419
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 411 -----PPLISWKAQETVPMNIIILGGGFAMAKGCEESGLSAWIGQLHLEHV 465
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 420 VADYSEPLITWKEFQSFMPMDIALVGGGFALADGCEESGLSKWIGKSLPGLPAWLI 479
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 466 VLLITVIAFFTEFASNTATIIIFLPVLAELAIRLHVHPLYLMIPTVSCSYAFM 525
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 480 ILISLWMTISLTVASNPNATITIFLPILSPLABAIHVNPILYIIPSTCLTSFAPL 539
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 526 PKNSTAFSTGHLVKKDMVTGLLMNMVGLLSLMMNTWAQAIIFQLGTFPDWANT 585
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```


Db 540 PPAIVPSYGLKVIDMWKAGLVNIVGAVVWMLGICTWIVPFDLYTYPWSA-----592
QY 586 TALPPALTNNTV 597
Db 593 ----PAMSNETM 600

RESULT 3

US-09-556-916-2

; Sequence 2, Application US/09556916

; Patent No. 6548271

; GENERAL INFORMATION:

; APPLICANT: Turner, Alex

; APPLICANT: Zambrowicz, Brian

; APPLICANT: Nehls, Michael

; APPLICANT: Friedrich, Glenn

; APPLICANT: Sands, Arthur T.

; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins

; FILE REFERENCE: 8535-041-999

; CURRENT APPLICATION NUMBER: US/09/556,916

; CURRENT FILING DATE: 2000-04-21

; NUMBER OF SEQ ID NOS: 32

; SOFTWARE: FaastSeq for Windows Version 3.0

; SEQ ID NO 2

; LENGTH: 627

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-556-916-2

Query Match 38.4%; Score 1188; DB 4; Length 627;

Best Local Similarity 38.6%; Pred. No. 5e-111;

Matches 241; Conservative 127; Mismatches 194; Indels 62; Gaps 12;

QY 15 RRLVLLVLPALLPILFALPPKGRCLYVILLMAVYVWCTEALPLSVTALLPIILPPMG 74
Db 11 RKLWVCPVLLPVLHPSEASCAVILVAVYVWSEAVPLGAALVPAFLYPPFG 70

QY 75 ILPSSKVCQYFLDTNLFSLGLMASAIEERNLHRRALKVLMVGVQPARLIIGMMVT 134
Db 71 VLRSNEVAAYEFKNTLLLVGVICVAAAVEKNLHRRALKVLMVGVQPARLIIGMMVT 130

QY 135 TSFLSWLNTASTAMMLPIASAILKSLF-----GQDTRKDLF-----173
Db 131 TLLSWLNTSTTAWMPIVEAVLQELVSADEQLVAGNSNTEAEPISLDVKNQSPSL 190

QY 174 -----REGDSTAARVGNGLRTPV--TEMQFLASSEGG---HAEDVEAPLELPDDSK 220
Db 191 ELIFVNEEDRSNADLTLLMHNENLNGVPSITNPIKTANQHOGKKQHPQVLTSPRK 250

QY 221 EE-----EHRNINWKGFLISIPYSASIGGTATLTGTAPNLIILGOLKSPFPQCDVNF 273
Db 251 QKLNRYRSHHDQMKCLSLISYSATIGLTTIIGTSTLSIFLEHFNNOYPAAEVNF 310

QY 274 GSWFIFAPPLMLLLVGLWLTSLFYLGMSWRGW----RKKNKSLQDVAEDKAKAVIQEE 329
Db 311 GTWFLFSPISLIMLVVSWFMWHLPLGCFNFKETCSLSKKKTKREQLSEKR----IQEE 366

QY 330 FQNLGPIKFAEQAVFLFCLFALLFSRDPKPIGWASLIF-APGFVSDAVTGVAVITLF 388
Db 367 YEKLGDISYPMVMTGFFILMTVLMFTREPGVPGWDSFFKEKKGYRTDATVSFVLGFLF 426

QY 389 FPPSQKSLKWWDFDK--APNSE-----TEPLLSWKKAEQETVPWNIIILGGGFAMAKGE 442
Db 427 LIPAKKPC-----FGKNDGENQEHSLGTEPIITWKDFQKTMPEWIVLVGGYALASGK 482

QY 443 ESGLSAWIGQHLPLEHVPPLAVLLITVIAFFTEPASNTATIIIFLPVLAELAIRLHV 502
Db 483 SSGLSWTIGNQMLSSSLPPWAVTLACILVSVTEFVSNPATITIFILPILCSLSETHI 542

QY 503 HPLYLMIPGTVSCSYAFMLPVSTPPNSIAFTGHLLVKDMVTRTGLLMNLMGVLLSLAMN 562
Db 543 NPLYTLIPVTMCISFAVMLPVGNPPNAIVFSYGHQIKDMVKAGLVNIVGLVIMVAIN 602

QY 563 TWAQAIQFOLGTTPDPWANTHAANVT 586
Db 603 TWGVSFLFHLTYPAWA--RVSNIT 624

RESULT 4

US-09-949-016-6840

; Sequence 6840, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FaastSeq for Windows Version 4.0

; SEQ ID NO 6840

; LENGTH: 627

; TYPE: PRT

; ORGANISM: Human

US-09-949-016-6840

Query Match 38.3%; Score 1186; DB 4; Length 627;

Best Local Similarity 38.5%; Pred. No. 8e-111;

Matches 240; Conservative 128; Mismatches 194; Indels 62; Gaps 12;

QY 15 RRLVLLVLPALLPILFALPPKGRCLYVILLMAVYVWCTEALPLSVTALLPIILPPMG 74
Db 11 RKLWVCPVLLPVLHPSEASCAVILVAVYVWSEAVPLGAALVPAFLYPPFG 70

QY 75 ILPSSKVCQYFLDTNLFSLGLMASAIEERNLHRRALKVLMVGVQPARLIIGMMVT 134
Db 71 VLRSNEVAAYEFKNTLLLVGVICVAAAVEKNLHRRALKVLMVGVQPARLIIGMMVT 130

QY 135 TSFLSWLNTASTAMMLPIASAILKSLF-----GQDTRKDLF-----173
Db 131 TLLSWLNTSTTAWMPIVEAVLQELVSADEQLVAGNSNTEAEPISLDVKNQSPSL 190

QY 174 -----REGDSTAARVGNGLRTPV--TEMQFLASSEGG---HAEDVEAPLELPDDSK 220
Db 191 ELIFVNEEDRSNADLTLLMHNENLNGVPSITNPIKTANQHOGKKQHPQVLTSPRK 250

QY 221 EE-----EHRNINWKGFLISIPYSASIGGTATLTGTAPNLIILGOLKSPFPQCDVNF 273
Db 251 QKLNRYRSHHDQMKCLSLISYSATIGLTTIIGTSTLSIFLEHFNNOYPAAEVNF 310

QY 274 GSWFIFAPPLMLLLVGLWLTSLFYLGMSWRGW----RKKNKSLQDVAEDKAKAVIQEE 329
Db 311 GTWFLFSPISLIMLVVSWFMWHLPLGCFNFKETCSLSKKKTKREQLSEKR----IQEE 366

QY 330 FQNLGPIKFAEQAVFLFCLFALLFSRDPKPIGWASLIF-APGFVSDAVTGVAVITLF 388
Db 367 YEKLGDISYPMVMTGFFILMTVLMFTREPGVPGWDSFFKEKKGYRTDATVSFVLGFLF 426

QY 389 FPPSQKSLKWWDFDK--APNSE-----TEPLLSWKKAEQETVPWNIIILGGGFAMAKGE 442
Db 427 LIPAKKPC-----FGKNDGENQEHSLGTEPIITWKDFQKTMPEWIVLVGGYALASGK 482

QY 443 ESGLSAWIGQHLPLEHVPPLAVLLITVIAFFTEPASNTATIIIFLPVLAELAIRLHV 502
Db 483 SSGLSWTIGNQMLSSSLPPWAVTLACILVSVTEFVSNPATITIFILPILCSLSETHI 542

QY 503 HPLYLMIPGTVSCSYAFMLPVSTPPNSIAFTGHLLVKDMVTRTGLLMNLMGVLLSLAMN 562
Db 543 NPLYTLIPVTMCISFAVMLPVGNPPNAIVFSYGHQIKDMVKAGLVNIVGLVIMVAIN 602

```
QY 563 TWAQAFQLGTFPDWANTHAANT 586
Db 603 TWGVSFLHLDTPAWA--RVSNT 624

RESULT 5
US-09-556-916-14
; Sequence 14, Application US/09556916
; Patent No. 6548271
; GENERAL INFORMATION:
; APPLICANT: Turner, Alex
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins
; FILE REFERENCE: 8535-041-999
; CURRENT APPLICATION NUMBER: US/09/556,916
; CURRENT FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-556-916-14

Query Match 38.3%; Score 1183.5; DB 4; Length 626;
Best Local Similarity 38.5%; Pred. No. 1.4e-110;
Matches 240; Conservative 127; Mismatches 195; Indels 61; Gaps 12;

QY 15 RRLVLLVPLALPILFALPPKRGCLYVILLMAVYVCTEALPLSVTALLPILPPFMG 74
Db 11 RKLIVVCPVPLLLPLPVLHPSSSEASCAYVLIIVTAVYVSEAVPLGAAALVPAFLYPPFG 70
QY 75 ILPSKVCPOYFLDTNFIPLSLGIMASAEIRNLHRIALKVLMLVGVOPARILGMVT 134
Db 71 VLRSNEVAEEFQNTLLLVGVICVAAAEKWNHRIALRMVLMAKAGPGLMLLCFMC 130
QY 135 TSFLSMWLSNTASTAMPLPIASAILKSLF-----GORDTRKDL----- 173
Db 131 TTLSSMWLSNTSTTAWMPIVEAVLQELVSAEDQVAGNSNTEEAEPISLDVKNQSPSL 190
QY 174 -----REGEDSTAARVGNGLRTP--TEMQFLASSEGG---HAEDVEAPLELPDDSK 221
Db 191 ELIFVNEESNADLTLMHNENLNGVPSITNPITKANQHQKQHPSEKQPVLTPSPRKQ 250
QY 222 E-----EHRNINWKGLFISIPYSASIGGTATLTGTAPNLILGLKSPFPQCDVVNFG 274
Db 251 KLNKRYRSHDDQMKCKLSLSISYSATIGLTTIIGTSTLSIFLEHNNQYPAAEVNFG 310
QY 275 SWFIAPFPLMLLFLLVGWLMTISFLYGGMSWRGW----RKNSKLQDVADKAKAVIQEEF 330
Db 311 TWFLSPFISLIMLVSWFWHMLFLGCKPKETCSLSKKKTKREQLSEK-----IQEEY 366
QY 331 QNLGPIKFAEQAVFILCLFAILLFSDPKFIPGWASLF-APGFVSDAVTGVAIVTILFF 389
Db 367 EKLGDISYPEMVTGFFFTILMTVLTREPGVPGWDSFFEKKGYRTDATVSVFLGFLFL 426
QY 390 FPSQKPSLKNWDFK--APNSE-----TEPLLSSWKAQETVPWNIIILLGGGFAMAKGCE 443
Db 427 IPAKKPC-----FGKNNGENQESLGTETPIITWKDFQKTPWEIVLVGGYALASGKS 482
QY 444 SGLSAWIGQLHPLHVPPLLAVALITVIAFFTEPASNTATITIFLPVLAELAIRLHV 503
Db 483 SGLSTWIGNQMLSLSSLPWAVTLLACILVSVTEFVSNPATITIFLPILCSLSETHIN 542
QY 504 PLYLMPGTVSCSAFMLPVSTPNSTAFSTGHLLVVKDMVKTGLLMLMGVLLLSLAMNT 563
Db 543 PLYTLIPVTMCISFAVMLPVGNPPNAIVFSYGHQCIKDMVKAGLVNVLGLVIMVAINT 602
QY 564 WQAQIFQLGTFPDWANTHAANT 586
```

```
Db 603 WGVSLFHLDTYPAWA--RVSNT 623

RESULT 6
US-09-556-916-4
; Sequence 4, Application US/09556916
; Patent No. 6548271
; GENERAL INFORMATION:
; APPLICANT: Turner, Alex
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins
; FILE REFERENCE: 8535-041-999
; CURRENT APPLICATION NUMBER: US/09/556,916
; CURRENT FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 627
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-556-916-4

Query Match 38.2%; Score 1182; DB 4; Length 627;
Best Local Similarity 38.5%; Pred. No. 2e-110;
Matches 240; Conservative 127; Mismatches 195; Indels 62; Gaps 12;

QY 15 RRLVLLVPLALPILFALPPKRGCLYVILLMAVYVCTEALPLSVTALLPILPPFMG 74
Db 11 RKLIVVCPVPLLLPLPVLHPSSSEASCAYVLIIVTAVYVSEAVPLGAAALVPAFLYPPFG 70
QY 75 ILPSKVCPOYFLDTNFIPLSLGIMASAEIRNLHRIALKVLMLVGVOPARILGMVT 134
Db 71 VLRSNEVAEEFQNTLLLVGVICVAAAEKWNHRIALRMVLMAKAGPGLMLLCFMC 130
QY 135 TSFLSMWLSNTASTAMPLPIASAILKSLF-----GORDTRKDL----- 173
Db 131 TTLSSMWLSNTSTTAWMPIVEAVLQELVSAEDQVAGNSNTEEAEPISLDVKNQSPSL 190
QY 174 -----REGEDSTAARVGNGLRTP--TEMQFLASSEGG---HAEDVEAPLELPDDSK 220
Db 191 ELIFVNEESNADLTLMHNENLNGVPSITNPITKANQHQKQHPSEKQPVLTPSPRK 250
QY 221 E-----EHRNINWKGLFISIPYSASIGGTATLTGTAPNLILGLKSPFPQCDVVNFG 273
Db 251 QKLNKRYRSHDDQMKCKLSLSISYSATIGLTTIIGTSTLSIFLEHNNQYPAAEVNFG 310
QY 274 GSWFIAPFPLMLLFLLVGWLMTISFLYGGMSWRGW----RKNSKLQDVADKAKAVIQEE 329
Db 311 GTWFLSPFISLIMLVSWFWHMLFLGCKPKETCSLSKKKTKREQLSEK-----IQEE 366
QY 330 QNLGPIKFAEQAVFILCLFAILLFSDPKFIPGWASLF-APGFVSDAVTGVAIVTILF 388
Db 367 YEKLGDISYPEMVTGFFFTILMTVLTREPGVPGWDSFFEKKGYRTDATVSVFLGFLFL 426
QY 389 FPSQKPSLKNWDFK--APNSE-----TEPLLSSWKAQETVPWNIIILLGGGFAMAKGCE 442
Db 427 LIPAKKPC-----FGKNNGENQESLGTETPIITWKDFQKTPWEIVLVGGYALASGSK 482
QY 443 ESSLASWIGQLHPLHVPPLLAVALITVIAFFTEPASNTATITIFLPVLAELAIRLHV 502
Db 483 SGLSTWIGNQMLSLSSLPWAVTLLACILVSVTEFVSNPATITIFLPILCSLSETHI 542
QY 503 HPLYLMPGTVSCSAFMLPVSTPNSTAFSTGHLLVVKDMVKTGLLMLMGVLLLSLAMN 562
Db 543 NPLYTLIPVTMCISFAVMLPVGNPPNAIVFSYGHQCIKDMVKAGLVNVLGLVIMVAIN 602
QY 563 TWAQAFQLGTFPDWANTHAANT 586
Db 603 TWGVSFLHLDTPAWA--RVSNT 624
```

RESULT 8

US-09-556-916-16
; Sequence 16, Application US/09556916
; Patent No. 6548271
; GENERAL INFORMATION:
; APPLICANT: Turner, Alex
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins
; FILE REFERENCE: 8535-041-999
; CURRENT APPLICATION NUMBER: US/09/556,916
; CURRENT FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-556-916-16

Query Match 38.1%; Score 1177.5; DB 4; Length 626;
Best Local Similarity 38.4%; Pred. No. 5.8e-110;
Matches 239; Conservative 127; Mismatches 196; Indels 61; Gaps 12;
QY 15 RRLVLLVPLALLPILFALPKGRCGLVILLMAVYVWCTEALPLSVTALLPILPFPFG 74
Db 11 RKLLVVCVPLLLPLVLPVHPSEASCAYLIVAVYVSEAVPLGAALVPAFLYPPFG 70
QY 75 ILPSKVCQYFLDTNFLFLSGLIMASAEERNLHRRIALKVLMLVGVQPARLILGMVT 134
Db 71 VLRSNEVAEYFNKNTLLVGVICVAAAEKWNHKLRIALRMVLMAGAKPGMLLCPMCC 130
QY 135 TSFLSMWLSNTASTAMMLPIASAILKSLF-----GQDTRKDL- 173
Db 131 TLLSMLWLSNTSTTAMVPIVEAVLQELVSADEQLVAGNSNTEEAEPISLDVKNQSPSL 190
QY 174 -----REGDSTAARVNGRLTVP--TEMQFLASSEGG--HAEDVEAPLELPDDSK 221
Db 191 ELIFVNEBSNADLTLLMHNENLNGVPSITNPIKTANQHQKKQHPQKQVLTSPRKQ 250
QY 222 E-----EHRNINWKGLISIPYSASIGGTATLTGTAPNLIILGQLKSPFPQCDVNF 274
Db 251 KLNRYRSHHDQMICCKLSLSISYSATIGLTTIGTSTLSIFLHFNNQYPAAEVNF 310
QY 275 GSWFIPAPLMMLLLVGLWLSFLYVGMWGRW-----RKNSKLQDVADKAKAVIQEF 330
Db 311 GTWFLFSPISLIMLVSWFMWHLFLGCFKTCSLSKKKTKREQLSEKR---IQEE 366
QY 331 QNLGPIKFAEQAVFLFCLFAILLFSRDPKPIPGWASLF-APGFVSDAVTGVAIVTILF 389
Db 367 EKLGDISYPMVMTGFFFLMTVLMFTREPGVPGWDSFEKKGYRTDATSVFGLFLF 426
QY 390 FPSQKPSLKNWFDK--APNSE-----TEPLLSWKAQETVPWNILILGGGFAMAKGCE 443
Db 427 LIPAKKPC---FGKNDGENQHSLSGTESIITWKDFQKTMPEWIVILVGGYALASGSK 482
QY 444 ESSLGSAWIGQQLHPLHVPPLAVLLITVVIAPFTFASNTATIIIFLPVLAELAIRLHV 503
Db 483 SGLTWTIGNQMLSSLPWAVTLLACILVSIVTEFVSNPATITIFLPILCSLSETLHI 542
QY 504 PLYLMIPTVSCSAFVFLPVSTPNSTAFSGHLLVKDMVTRTGLLMLMGVLLSLANNT 563
Db 543 NPLYTLIPVTMCISFAVMLPVGNPNNAIVFSYGHQCIKDMVKAGLVNIGLVIMVAIN 602
QY 564 WQAQIFOLGTFPPDWTANHAANVT 586
Db 603 TWGVSFLHLDITYPAWA--RVSNIT 624

RESULT 9

US-09-556-916-20

RESULT 7

US-09-556-916-8
; Sequence 8, Application US/09556916
; Patent No. 6548271
; GENERAL INFORMATION:
; APPLICANT: Turner, Alex
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins
; FILE REFERENCE: 8535-041-999
; CURRENT APPLICATION NUMBER: US/09/556,916
; CURRENT FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 627
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-556-916-8

Query Match 38.1%; Score 1180; DB 4; Length 627;
Best Local Similarity 38.5%; Pred. No. 3.2e-110;
Matches 240; Conservative 127; Mismatches 195; Indels 62; Gaps 12;
QY 15 RRLVLLVPLALLPILFALPKGRCGLVILLMAVYVWCTEALPLSVTALLPILPFPFG 74
Db 11 RKLLVVCVPLLLPLVLPVHPSEASCAYLIVAVYVSEAVPLGAALVPAFLYPPFG 70
QY 75 ILPSKVCQYFLDTNFLFLSGLIMASAEERNLHRRIALKVLMLVGVQPARLILGMVT 134
Db 71 VLRSNEVAEYFNKNTLLVGVICVAAAEKWNHKLRIALRMVLMAGAKPGMLLCPMCC 130
QY 135 TSFLSMWLSNTASTAMMLPIASAILKSLF-----GQDTRKDL- 173
Db 131 TLLSMLWLSNTSTTAMVPIVEAVLQELVSADEQLVAGNSNTEEAEPISLDVKNQSPSL 190
QY 174 -----REGDSTAARVNGRLTVP--TEMQFLASSEGG--HAEDVEAPLELPDDSK 220
Db 191 ELIFVNEBSNADLTLLMHNENLNGVPSITNPIKTANQHQKKQHPQKQVLTSPRKQ 250
QY 221 EE-----EHRNINWKGLISIPYSASIGGTATLTGTAPNLIILGQLKSPFPQCDVNF 273
Db 251 KLNRYRSHHDQMICCKLSLSISYSATIGLTTIGTSTLSIFLHFNNQYPAAEVNF 310
QY 274 GSWFIPAPLMMLLLVGLWLSFLYVGMWGRW-----RKNSKLQDVADKAKAVIQEF 329
Db 311 GTWFLFSPISLIMLVSWFMWHLFLGCFKTCSLSKKKTKREQLSEKR---IQEE 366
QY 330 QNLGPIKFAEQAVFLFCLFAILLFSRDPKPIPGWASLF-APGFVSDAVTGVAIVTILF 388
Db 367 EKLGDISYPMVMTGFFFLMTVLMFTREPGVPGWDSFEKKGYRTDATSVFGLFLF 426
QY 389 FPSQKPSLKNWFDK--APNSE-----TEPLLSWKAQETVPWNILILGGGFAMAKGCE 442
Db 427 LIPAKKPC---FGKNDGENQHSLSGTESIITWKDFQKTMPEWIVILVGGYALASGSK 482
QY 443 ESSLGSAWIGQQLHPLHVPPLAVLLITVVIAPFTFASNTATIIIFLPVLAELAIRLHV 502
Db 483 SGLTWTIGNQMLSSLPWAVTLLACILVSIVTEFVSNPATITIFLPILCSLSETLHI 542
QY 503 PLYLMIPTVSCSAFVFLPVSTPNSTAFSGHLLVKDMVTRTGLLMLMGVLLSLANNT 562
Db 543 NPLYTLIPVTMCISFAVMLPVGNPNNAIVFSYGHQCIKDMVKAGLVNIGLVIMVAIN 602
QY 563 WQAQIFOLGTFPPDWTANHAANVT 586
Db 603 TWGVSFLHLDITYPAWA--RVSNIT 624

; Sequence 20, Application US/09556916
; Patent No. 6548271

; GENERAL INFORMATION:
; APPLICANT: Turner, Alex
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Sands, Arthur T.

; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins

; FILE REFERENCE: 8535-041-999

; CURRENT APPLICATION NUMBER: US/09/556,916

; CURRENT FILING DATE: 2000-04-21

; NUMBER OF SEQ ID NOS: 32

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 20

; LENGTH: 626

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-556-916-20

Query Match 38.0%; Score 1175.5; DB 4; Length 626;

Best Local Similarity 38.4%; Pred. No. 9.2e-110;

Matches 239; Conservative 127; Mismatches 196; Indels 61; Gaps 12;

```
QY 15 RRLVLLVPLALLPLFALPPKRGCLYVILLMAVYWCTEALPLSVTALLPIILPFPMG 74
Db 11 RKLLVVCVPLLLPLPLVLPSPSEASCAVLIIVTAVYVSEAVPLGAAALVPALFPPFG 70
QY 75 ILPSSKVCQVFLDTNLFSLGLIMASAJEERNLHRRIALKVLMLGVQPARILGMVMT 134
Db 71 VLRSNEVAAYEYFNKNTLLLVGVICVAAAVEKNLHKRIALRMVLMAGAKPGMLLCFMC 130
QY 135 TSFLSMWLSNTASTAMMLPIASAILKSLF-----GQDTRKDLP----- 173
Db 131 TTLLSMWLSNTSTTAMVMPIVEAVLOELVSAEDEQLVAGNSNTEEAEPISLDVKNQSPSL 190
QY 174 -----REGEDSTAAVRGNGLRTP--TEMQFLASSEGG---HAEDVEAPLELPDPSKE 221
Db 191 ELIFVNESNADLTTLMHENENLNGVPSITNPITANOHQKQHPSEKQVLTSPRKQ 250
QY 222 E-----EHRNRNWKGLFISIPYSASIGGTATLTGTAPNLLILGOLKSPFQCDVVNFG 274
Db 251 KLNKRYRSHHDQMICKLSLSISYSATIGLTIIGTSTLSLIFLEHFNNOYPAAEVNFG 310
QY 275 SWTFAPAPMLLFLVGLWLTISFLYCGMSWRGW----RKNSKLDQVAEDKAKAVIOEEF 330
Db 311 TWFLFSPFISLIMLVSWFMWHLFLGNCFKETCSLSKKKKTKRQLSEKR-----IQEE 366
QY 331 QNLGPIKFAEQVFLFCLFALLFSRDPKPIPGWASLF-APGFVSDAVTGVAVITILFF 389
Db 367 EKLGDISYPEWVTGFFILMTVLFTRFPGVPGWDSFFEKKGYRTDATVSVFLGFLFL 426
QY 390 FPSQKPSLKWWDFFK--APNSE-----TEPLLSSWKAQETVPMNILLGGGFAMAKGCEE 443
Db 427 IPAKKPC----FGKNDGENQHSLSGTSIITWKDFQKTPMPEWIVILVGGGYALASGSKS 482
QY 444 SGLSAMIQGLHPLHVPPLAVLLITVIAFFTEFASNTATIIIFLPVLAELAIRLHVH 503
Db 483 SGLSTWIGNQMLSSLSLPPMAVTLACILSVITFVSNPATITIFLPILCSLSETHIN 542
QY 504 PLYLMIPTVSCSYAFMLPVSTPNSTAFSTGHLLVKDMVRTGILLMNLGMVLLSLAMNT 563
Db 543 PLYTLIPVTMCISFAVMLPVGNPNNAIVFSYGHQIKDMVKAGLVGNVIGLVIMVAINT 602
QY 564 WQAQIFOLGTFPPDWANTHAANVT 586
Db 603 WGVSLFHLDTYPAWA--RVSNTIT 623
```

RESULT 10

US-09-556-916-10

; Sequence 10, Application US/09556916

; Patent No. 6548271

; APPLICANT: Turner, Alex

RESULT 11

US-09-556-916-22

; Sequence 22, Application US/09556916

; Patent No. 6548271

; GENERAL INFORMATION:

; APPLICANT: Turner, Alex

; GENERAL INFORMATION:

; APPLICANT: Turner, Alex

; APPLICANT: Zambrowicz, Brian

; APPLICANT: Nehls, Michael

; APPLICANT: Friedrich, Glenn

; APPLICANT: Sands, Arthur T.

; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins

; FILE REFERENCE: 8535-041-999

; CURRENT APPLICATION NUMBER: US/09/556,916

; CURRENT FILING DATE: 2000-04-21

; NUMBER OF SEQ ID NOS: 32

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 10

; LENGTH: 627

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-556-916-10

Query Match 37.9%; Score 1174; DB 4; Length 627;

Best Local Similarity 38.3%; Pred. No. 1.3e-109;

Matches 239; Conservative 127; Mismatches 196; Indels 62; Gaps 12;

```
QY 15 RRLVLLVPLALLPLFALPPKRGCLYVILLMAVYWCTEALPLSVTALLPIILPFPMG 74
Db 11 RKLLVVCVPLLLPLPLVLPSPSEASCAVLIIVTAVYVSEAVPLGAAALVPALFPPFG 70
QY 75 ILPSSKVCQVFLDTNLFSLGLIMASAJEERNLHRRIALKVLMLGVQPARILGMVMT 134
Db 71 VLRSNEVAAYEYFNKNTLLLVGVICVAAAVEKNLHKRIALRMVLMAGAKPGMLLCFMC 130
QY 135 TSFLSMWLSNTASTAMMLPIASAILKSLF-----GQDTRKDLP----- 173
Db 131 TTLLSMWLSNTSTTAMVMPIVEAVLOELVSAEDEQLVAGNSNTEEAEPISLDVKNQSPSL 190
QY 174 -----REGEDSTAAVRGNGLRTP--TEMQFLASSEGG---HAEDVEAPLELPDPSK 220
Db 191 ELIFVNESNADLTTLMHENENLNGVPSITNPITANOHQKQHPSEKQVLTSPRK 250
QY 221 EE-----EHRNRNWKGLFISIPYSASIGGTATLTGTAPNLLILGOLKSPFQCDVVNFG 273
Db 251 QLNKRYRSHHDQMICKLSLSISYSATIGLTIIGTSTLSLIFLEHFNNOYPAAEVNFG 310
QY 274 GSWFIPAPMLLFLVGLWLTISFLYCGMSWRGW----RKNSKLDQVAEDKAKAVIOEE 329
Db 311 GTWFLFSPFISLIMLVSWFMWHLFLGNCFKETCSLSKKKKTKRQLSEKR-----IQEE 366
QY 330 QNLGPIKFAEQVFLFCLFALLFSRDPKPIPGWASLF-APGFVSDAVTGVAVITILF 388
Db 367 YEKLGDISYPEWVTGFFILMTVLFTRFPGVPGWDSFFEKKGYRTDATVSVFLGFLFL 426
QY 389 FPSQKPSLKWWDFFK--APNSE-----TEPLLSSWKAQETVPMNILLGGGFAMAKGCE 442
Db 427 LIPAKKPC----FGKNDGENQHSLSGTSIITWKDFQKTPMPEWIVILVGGGYALASGSK 482
QY 443 ESSLAWIGQGLHPLHVPPLAVLLITVIAFFTEFASNTATIIIFLPVLAELAIRLHV 502
Db 483 SGLSTWIGNQMLSSLSLPPMAVTLACILSVITFVSNPATITIFLPILCSLSETHI 542
QY 503 HPLYLMIPTVSCSYAFMLPVSTPNSTAFSTGHLLVKDMVRTGILLMNLGMVLLSLAMN 562
Db 543 NPLYTLIPVTMCISFAVMLPVGNPNNAIVFSYGHQIKDMVKAGLVGNVIGLVIMVAIN 602
QY 563 WQAQIFOLGTFPPDWANTHAANVT 586
Db 603 TWGVSLFHLDTYPAWA--RVSNTIT 624
```



```

; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins
;
; FILE REFERENCE: 8535-041-999
;
; CURRENT APPLICATION NUMBER: US/09/556,916
;
; CURRENT FILING DATE: 2000-04-21
;
; NUMBER OF SEQ ID NOS: 32
;
; SOFTWARE: FastSeq for Windows Version 3.0
;
; SEQ ID NO 18
;
; LENGTH: 580
;
; TYPE: PRT
;
; ORGANISM: Homo sapiens
;
US-09-556-916-18

```

	Query Match	32.8%	Score 1014.5	DB 4	Length 580
	Best Local Similarity	34.7%	Pred. No. 1.6e-93		
	Matches 216	Conservative 120	Mismatches 93	Indels 107	Gaps 14
Qy	15	RLLLVLLVPLALPLIPALPPKEGRCLXYVLLMAVYCTEALPLSVTALLPIILFFPMG	74		
Db	11	RKLLVVCVPLLLPLPVLPHPSSRASCAYLVITAVYVSEAVPLGAALVPALYFPFG	70		
Qy	75	ILPSSKVCPOYFLDTNFLFSLGLMASAIEERNLHRIALKVLMLVGVPARLLILGMVMT	134		
Db	71	VLRSEVAAEYFKNTLLLVGVICVAAAEKWNLHKRIALRMVLMAGAKPGMLLLCFMCC	130		
Qy	135	TSFLSMLNSNTASTAMMLPIASAILKSLF-----GQDRTKDLPL-----	173		
Db	131	TTLSLMLNSITSTAMVMPIVEAVLQELVSAEDSQLVAGNSNTTEAPEISLDVKNSPSL	190		
Qy	174	-----REGEDSTAARVGNGLRTP--TEMQFLASSEGG---HAEDVEAPLELPDDSK	221		
Db	191	ELIPVNEESNADLTTLTHNENLNGVPSITNPDKTANQHQKQHPQSQEKQVLTSPSRKQ	250		
Qy	222	E-----EHRNINWKGFLISIPYSASIGTATLTGTAPNLLILGOLKSLFFPQCDVNVFG	274		
Db	251	KLNRKYRSHDQMCCKLSLSISYATIGGUTTIIGTSTSLIFLHFNQVPAAEVNVFG	310		
Qy	275	SWFIAPPLMLLFLVGLWLTISFYLGMSWRGW---RKNKSKLDQVDAEDKAKAVIOEEF	330		
Db	311	TWFLFSPISILMLVSVFWHMLFLGCNPKETCSLKKKTKREQLSEK----IQEY	366		
Qy	331	QNLGPKEAQVAILFCFLFAILLFSRDPKPIPGWASLF-APGFVSDAVTGAIVTILFF	389		
Db	367	EKLGDISYPENVMTGFFFTILMTVLMTFTREPGVPGWDSFEKKGYKTDATSVFJGLLFL	426		
Qy	390	PPSOKPSLKNWDFK--APNSE-----TEPLLSWKKAQETVPWNIILLGGGFAMAKGEE	443		
Db	427	IPAKPC-----FGKNQDGENQHSIGTEPIITWKDFQKTMPEIIVILVGGGVALASGKS	482		
Qy	444	SGLSAGTGGQLHPIEHVDPPLAVLLITVVIATFTFASNTATITIFLPVLAELAIRLVH	503		
Db	483	SGLSTWIGNQMLSSLPWPVATLLACILVSIIVTFEVSNPATITIFLPI-----	532		
Qy	504	PLYLMIPTGVSCSYAFMLPVSTPPNSIAFSTGHLLLVKDMVTRTGLMLNMGVLLLSLAMNT	563		
Db	533	-----CS-----LVKAGLVNVGLVIVMVAINT	556		
Qy	564	WQAQIFQLGTFPDWNTHAANVT	586		
Db	557	WGVSLFHLDTYPAWA--RVSNIT	577		

RESULT 14
US-09-556-916-12
; Sequence 12, Application US/09556916
; Patent No. 6548271
; GENERAL INFORMATION:
; APPLICANT: Turner, Alex
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6548271el
; FILE REFERENCE: 8535-041-999

```

; CURRENT APPLICATION NUMBER: US/09/556,916
; CURRENT FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 581
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-556-916-12

```

Query Match	32.78;	Score 1011;	DB 4;	Length 581;
Best Local Similarity	34.69;	Pred. No. 3.5e-93;		
Matches 216;	Conservative 120;	Mismatches 180;	Indels 108;	Gaps 144
Qy	15	RLVLVLPLALPLILFALPKPEGRCLVYVLLMNAVYCTEALPLSVTALLPIILFFPMG	74	
Db	11	RKLLVVCVPLLLLPLVLPSPSEASCAVVLIVTAVYVWSEAVPLGAAALVPAFLYPPFG	70	
Qy	75	ILPSKVCQOYFLDTNPLFLSLGLIMASAIERNLHRRIALKVLMLGVOPARLLIGMMVT	134	
Db	71	VLRSNEVAABEYFKNTTLLLVGVICVAAAEVKNNLHKRIALRMVLIMAGAKPGMLLLCFMCC	130	
Qy	135	TSFLSMWLSNTASTAMMLPIATASATLKSLF-----GQDRTKDLPL-----	173	
Db	131	TTLSMWLSNTSTTAMVMPIVEAVLOELVSAEDQLVAGNSNTEEAEPISLDVKNQSPSL	190	
Qy	174	-----REGEDSTAARVGNGLTVP--TEMQFLASEGG---HAEDVEAPLELPDDSK	220	
Db	191	ELIPVNEDRSNADLTTLTMHNLNGLVPSITNPIQTANQHQKKQHPQOEKPQVLTSPRK	250	
Qy	221	EE-----EHRRNWKGFLISIPYSASIGGTATLTGTAPNLILLGOLKSPFQCDVNVF	273	
Db	251	QKLNKRYRSHHDMICKLSLSISYSAVIGLTTIIGTSTSLIFLEHNQYPAEVNVF	310	
Qy	274	GSWFIFAPPLMLLFLVGLWISFLYGGMSRWG-----RKNSKLQDVDAEDKAKAVIOEE	329	
Db	311	GTWFLFPISLIMLVVSWFMMHFLGCPKFCSTLSKKKTKREQLSEKR-----IQEE	366	
Qy	330	FQNLGPKFBAQVFIILFCLPAILLFRDPKPIQWASLF--APGFVSDAVTGVAVITLF	388	
Db	367	YEKLGDISYEMVVTGFFPILMTVLWFTRBQFVPGWDSFFBKKGYRTDTSVFLGFLF	426	
Qy	389	FPFQKPSLKWWDPK--APNSE-----TEPLLSKKQAEVTPVNIILLGGGFANAKGE	442	
Db	427	LIPAKKC-----FGKNDGENQEHSLGTESIITWKDFQKTWPWEIVLVGGYALASGK	482	
Qy	443	ESGLSANIGGQHPLEHVPPLLAVLLTVVIAFTPEPASNTATIIILFPLVLAELAIRLHV	502	
Db	483	SSGLSTWIGNQLSSLSLPPMAVTLACILSVITEFVSNPATITIFILPIL-----	533	
Qy	503	HPLYLMPGTVCSSYAFMLPVSTPPNSIAFTSGHLLVKDMVYRTGLMLMLMGVLLLSLANN	562	
Db	534	-----CS-----LVKAGLVGNVIGLIVVMVAIN	556	
Qy	563	TWQAIFQLGTFFPDWANTHAANVT	586	
Db	557	TWGVSLFHLDTYPAWA--RVSNIT	578	

RESULT 15
US-09-556-916-24
; Sequence 24, Application US/09556916
; Patent No. 6548271
; GENERAL INFORMATION:
; APPLICANT: Turner, Alex
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Sands, Arthur I.
; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins
; FILE REFERENCE: 8535-041-999
; CURRENT APPLICATION NUMBER: US/09/556, 916
; CURRENT FILING DATE: 2000-04-21

```
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 24
; LENGTH: 580
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-556-916-24

Query Match      32.5%; Score 1006.5; DB 4; Length 580;
Best Local Similarity 34.5%; Pred. No. 1e-92;
Matches 215; Conservative 120; Mismatches 181; Indels 107; Gaps 14;

QY 15 RLLVLLVPLALLPLFALPKEGECYVILLMAYVWCTEALPLSVTALLPIILPPMG 74
Db 11 RLLLVVCPVLLPLPLVLPSEASCAYLVITVAVYVSEAVPLGAAALVPAFLYPFG 70

QY 75 ILPSSKVCQYFLDTNFLSLGLIMASAEERNLHRIALKVLMLVGVOVARLIILGMVT 134
Db 71 VLRSNEVAAYEYFNVTLLLVGVICVAAVEKNLHRIALRMVLMAGAKPGMLLFCMCC 130

QY 135 TSFLSNWLSNTASTAMMLPIASAILKSLF-----GQDTRKDLF----- 173
Db 131 TTLLSNWLSNTSTAMVMPIVEAVLQELVSAEDEQLVAGNSNTEEAEPISLOVKNSQPSL 190

QY 174 -----REGEDSTAARNGLRTP--TEMQFLASSEG--HAEDVEAPLELPDDSKB 221
Db 191 ELIFVNEESNADLTILMHENENLNGVPSITNPITANHQKQKQPSQEKQPVLTPSPRKQ 250

QY 222 E-----EHRNINWKGFLISIPYASIGGTATLTCTAPNLILLGOLKSFPPQCDVNVFG 274
Db 251 KLNRYRSHHDQMICKUSLSISYATIGLTIIGTSTLIFLHFNNQYPAAEVNVFG 310

QY 275 SWFIAPFLMLFLVGLWGLWISFLYGGMSWRGW----RKNSKLQDVADKAKAVIQBEF 330
Db 311 TWFLSFPSILMLVSVNFWMHWFLGCFNFKETCSLSKKKTKRQLSEKR-----IQEY 366

QY 331 QNLGPIKFAEQAVFILFCLFAILLFSRDPKFTPGWASLP-APGFVSDAVTGVAIVTILFF 389
Db 367 EKLGDISYPEMVTGFFFIILMTVLFETREPFGVPGWDSFPEKKGRTDATSVSFLGLLFL 426

QY 390 FPSQPSLKWWFDFK--APNSE----TEPLLSWKKAOETVPWNIIILLGGGFAMAKGCEE 443
Db 427 IPAKKPC----FGKNDGENQEHSLGTSIITWKDFQKTPMWEIVILVGGGYALASGSKS 482

QY 444 SGLSAWIGQLHPLBHVPLLAFLITVVIAPFTEFASNTATIIIFLPVLAELAIRLHVH 503
Db 483 SGLSTWIGNQMLSSLSLPPWAVTLLACILVSVITEFVSNPATITIFLPI----- 532

QY 504 PLYLMIPGTVSCSYAFMLPVSTPPNSIAFSTGHLVKDMVRTGLLNMNLGMVLLSLAMNT 563
Db 533 -----CS-----LVKAGLGWVIGLVIVVVAINT 556

QY 564 WQAIFQLGTFPDWANTHAANVT 586
Db 557 WGVSLFLDTPAWA--RVSNT 577
```

Search completed: June 30, 2005, 09:02:19
Job time : 26.069 secs

THIS PAGE IS BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2005, 08:59:01 ; Search time 77.8579 Seconds
(without alignments)
2973.073 Million cell updates/sec

Title: US-10-017-479a-5

Perfect score: 3094

Sequence: 1 MAALAAAKVWSARRLLVL.....HAANYTALPPALTNNVTQL 600

Scoring table: BLASUM62

Gapop 10.0 , Gapext 0.5

Searched: 1722976 seqs, 385795295 residues

Total number of hits satisfying chosen parameters: 1722976

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3094	100.0	600	17	US-10-718-359-14
2	3092	99.9	600	14	US-10-167-994-13
3	2891	87.0	602	14	US-10-167-994-4
4	2326.5	75.2	520	15	US-10-108-260A-3862
5	1412.5	45.7	616	15	US-10-092-900A-46
6	1412.5	45.7	616	15	US-10-092-900A-48
7	1412.5	45.7	616	15	US-10-403-161-68
8	1411.5	45.6	568	13	US-10-173-519-2
9	1411.5	45.6	568	13	US-10-403-161-72
10	1411.5	45.6	568	16	US-10-490-080-1
11	1411.5	45.6	568	17	US-10-718-359-6
					Sequence 14, Appl
					Sequence 13, Appl
					Sequence 4, Appl
					Sequence 3862, Ap
					Sequence 46, Appl
					Sequence 48, Appl
					Sequence 68, Appl
					Sequence 2, Appl
					Sequence 7, Appl
					Sequence 1, Appl
					Sequence 6, Appl

Sequence 149, App
Sequence 4, Appl
Sequence 4, Appl
Sequence 44, Appl
Sequence 70, Appl
Sequence 12, Appl
Sequence 2, Appl
Sequence 2, Appl
Sequence 5, Appl
Sequence 12, Appl
Sequence 13, Appl
Sequence 6, Appl
Sequence 4, Appl
Sequence 23, Appl
Sequence 25, Appl
Sequence 10, Appl
Sequence 3, Appl
Sequence 11, Appl
Sequence 50, Appl
Sequence 74, Appl
Sequence 118, App
Sequence 5, Appl
Sequence 9, Appl
Sequence 2, Appl
Sequence 7, Appl
Sequence 14, Appl
Sequence 8, Appl
Sequence 16, Appl
Sequence 20, Appl
Sequence 10, Appl
Sequence 22, Appl
Sequence 52, Appl
Sequence 76, Appl
Sequence 6, Appl
Sequence 18, Appl
Sequence 12, Appl
Sequence 24, Appl
Sequence 2, Appl
Sequence 2, Appl
Sequence 8, Appl
Sequence 1947, Ap
Sequence 1949, Ap
Sequence 346, App
Sequence 10, Appl
Sequence 176, App
Sequence 436, App
Sequence 127443,
Sequence 160, App
Sequence 127444,
Sequence 516, App
Sequence 3754, App
Sequence 250656,
Sequence 22422, Ap
Sequence 254988,
Sequence 4, Appl
Sequence 354, App
Sequence 354, App
Sequence 354, App
Sequence 2634, Ap
Sequence 21334, A
Sequence 518, App
Sequence 8, Appl
Sequence 7, Appl
Sequence 1187, Ap
Sequence 594, App
Sequence 34252, A
Sequence 28038, A
Sequence 9458, Ap
Sequence 980, App

85 369 11.9 234 15 US-10-335-977-5397 Sequence 5397, Ap
86 366 11.8 453 15 US-10-369-493-9936 Sequence 9936, Ap
87 360 11.6 457 15 US-10-369-493-4266 Sequence 4266, Ap
88 352 11.4 236 16 US-10-767-701-43427 Sequence 43427, A
89 346 11.2 1006 17 US-10-741-849-7011 Sequence 7011, Ap
90 345 11.2 163 15 US-10-424-599-184881 Sequence 184881,
91 320 10.3 174 16 US-10-425-115-335649 Sequence 335649,
92 318.5 10.3 499 14 US-10-238-075-500 Sequence 500, App
93 308.5 10.0 300 16 US-10-767-701-46317 Sequence 46317, A
94 296 9.6 162 11 US-09-833-245-1345 Sequence 1945, Ap
95 295.5 9.6 162 11 US-09-833-245-1948 Sequence 1948, Ap
96 283 9.1 487 15 US-10-369-493-23559 Sequence 23559, A
97 263.5 8.5 164 15 US-10-424-599-205152 Sequence 205152,
98 252 8.1 54 9 US-09-864-761-46701 Sequence 46701, A
99 207.5 6.7 487 15 US-10-369-493-749 Sequence 749, App
100 203 6.6 109 16 US-10-767-701-44228 Sequence 44228, A

ALIGNMENTS

RESULT 1
US-10-718-359-14
; Sequence 14, Application US/10718359
; Publication No. US20050095240A1
; GENERAL INFORMATION:
; APPLICANT: MEDICAL COLLEGE OF GEORGIA RESEARCH INSTITUTE, INC.
; TITLE OF INVENTION: NACT AS A TARGET FOR LIFESPAN EXPANSION AND WEIGHT REDUCTION
; FILE REFERENCE: 275.0008 0101
; CURRENT APPLICATION NUMBER: US/10718,359
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/428,469
; PRIOR FILING DATE: 2002-11-22
; PRIOR APPLICATION NUMBER: 60/459,441
; PRIOR FILING DATE: 2003-04-01
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 14
; LENGTH: 600
; TYPE: PRT
; ORGANISM: rat NADC3
US-10-718-359-14

Query Match 100.0%; Score 3094; DB 17; Length 600;
Best Local Similarity 100.0%; Pred. No. 1.1e-273;
Matches 600; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAALAALAKKWSARRLLVLLVPLALLPILFALPPKGRCLYVILLMAVYWCTEALPLS 60
DB 1 MAALAALAKKWSARRLLVLLVPLALLPILFALPPKGRCLYVILLMAVYWCTEALPLS 60
QY 61 VTALLPILPPMGILPSSKVCPOYFLDTNPLFLSGLIMASAEERNLHRRALKVLMV 120
DB 61 VTALLPILPPMGILPSSKVCPOYFLDTNPLFLSGLIMASAEERNLHRRALKVLMV 120
QY 121 GVQPARLILGMVTTTSFLSMWLSNTASTAMMLPIASAILKSLFGQRTDKLPREGEDST 180
DB 121 GVQPARLILGMVTTTSFLSMWLSNTASTAMMLPIASAILKSLFGQRTDKLPREGEDST 180
QY 181 AAVRGNGLTRVPTQMFLASSEGGAEDVEAPLELPDDSKEEHRRNIWKGLISIPYSA 240
DB 181 AAVRGNGLTRVPTQMFLASSEGGAEDVEAPLELPDDSKEEHRRNIWKGLISIPYSA 240
QY 241 SIGGTATLTGTAPNLILGOLKSPFQCDVNVFGSWFIAPFLMLLLVGLWMLISFLYG 300
DB 241 SIGGTATLTGTAPNLILGOLKSPFQCDVNVFGSWFIAPFLMLLLVGLWMLISFLYG 300
QY 301 GMSWRGWRKNSKLQDVAEDKAKAVIQEFGNLGPIKFAEQAVFILFCLFAILLFSRDPK 360
DB 301 GMSWRGWRKNSKLQDVAEDKAKAVIQEFGNLGPIKFAEQAVFILFCLFAILLFSRDPK 360
QY 361 FIPGWASLFPAGFVSDAVTGVAVITLFFPFSQKPSLKWDFKAPNSETPEPLLSWKKAQ 420
DB 361 FIPGWASLFPAGFVSDAVTGVAVITLFFPFSQKPSLKWDFKAPNSETPEPLLSWKKAQ 420

DB 361 FIPGWASLFPAGFVSDAVTGVAVITLFFPFSQKPSLKWDFKAPNSETPEPLLSWKKAQ 420
QY 421 ETVPWNIIILLGGGFAMAKGCEESGLSAWIGQHLPHLHPVPLLAVALLIITVVIAPFTTEFA 480
DB 421 ETVPWNIIILLGGGFAMAKGCEESGLSAWIGQHLPHLHPVPLLAVALLIITVVIAPFTTEFA 480
QY 481 SNTATIIIFLPVLAELAIRLHVHPLYLMPGTVSCSYAFMLPVPSTPPNSIAFSTGHLLVK 540
DB 481 SNTATIIIFLPVLAELAIRLHVHPLYLMPGTVSCSYAFMLPVPSTPPNSIAFSTGHLLVK 540
QY 541 DMVRTGLMLNMGVLLLSLAMNTWAQAIQOLGTFPDWANTHANVTAALPPALTNNTVQTL 600
DB 541 DMVRTGLMLNMGVLLLSLAMNTWAQAIQOLGTFPDWANTHANVTAALPPALTNNTVQTL 600
RESULT 2
US-10-167-994-13
; Sequence 13, Application US/10167994
; Publication No. US20030082647A1
; GENERAL INFORMATION:
; APPLICANT: Reenan, Robert A.
; APPLICANT: Rogina, Blanka
; APPLICANT: Helfand, Stephen L.
; TITLE OF INVENTION: TRANSPORTER PROTEIN
; FILE REFERENCE: 13407-013001
; CURRENT APPLICATION NUMBER: US/10/167,994
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/255,013
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 10/017,479
; PRIOR FILING DATE: 2001-12-12
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 600
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-167-994-13

Query Match 99.9%; Score 3092; DB 14; Length 600;
Best Local Similarity 99.8%; Pred. No. 1.6e-273;
Matches 599; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAALAALAKKWSARRLLVLLVPLALLPILFALPPKGRCLYVILLMAVYWCTEALPLS 60
DB 1 MAALAALAKKWSARRLLVLLVPLALLPILFALPPKGRCLYVILLMAVYWCTEALPLS 60
QY 61 VTALLPILPPMGILPSSKVCPOYFLDTNPLFLSGLIMASAEERNLHRRALKVLMV 120
DB 61 VTALLPILPPMGILPSSKVCPOYFLDTNPLFLSGLIMASAEERNLHRRALKVLMV 120
QY 121 GVQPARLILGMVTTTSFLSMWLSNTASTAMMLPIASAILKSLFGQRTDKLPREGEDST 180
DB 121 GVQPARLILGMVTTTSFLSMWLSNTASTAMMLPIASAILKSLFGQRTDKLPREGEDST 180
QY 181 AAVRGNGLTRVPTQMFLASSEGGAEDVEAPLELPDDSKEEHRRNIWKGLISIPYSA 240
DB 181 AAVRGNGLTRVPTQMFLASSEGGAEDVEAPLELPDDSKEEHRRNIWKGLISIPYSA 240
QY 241 SIGGTATLTGTAPNLILGOLKSPFQCDVNVFGSWFIAPFLMLLLVGLWMLISFLYG 300
DB 241 SIGGTATLTGTAPNLILGOLKSPFQCDVNVFGSWFIAPFLMLLLVGLWMLISFLYG 300
QY 301 GMSWRGWRKNSKLQDVAEDKAKAVIQEFGNLGPIKFAEQAVFILFCLFAILLFSRDPK 360
DB 301 GMSWRGWRKNSKLQDVAEDKAKAVIQEFGNLGPIKFAEQAVFILFCLFAILLFSRDPK 360
QY 361 FIPGWASLFPAGFVSDAVTGVAVITLFFPFSQKPSLKWDFKAPNSETPEPLLSWKKAQ 420
DB 361 FIPGWASLFPAGFVSDAVTGVAVITLFFPFSQKPSLKWDFKAPNSETPEPLLSWKKAQ 420
QY 421 ETVPWNIIILLGGGFAMAKGCEESGLSAWIGQHLPHLHPVPLLAVALLIITVVIAPFTTEFA 480
DB 421 ETVPWNIIILLGGGFAMAKGCEESGLSAWIGQHLPHLHPVPLLAVALLIITVVIAPFTTEFA 480

Db 421 ETVPNWIIILLGGGFAMAKCESGLSAGWIGQLHPLHVPPLAVLLITVVIAFTTFA 480
QY 481 SNTATIIIFLPVLAELAIRLHVHPLYLMPGTVCSCYAFMLPVSTPPNSIAFSTGHLLVK 540
Db 481 SNTATIIIFLPVLAELAIRLHVHPLYLMPGTVCSCYAFMLPVSTPPNSIAFSTGHLLVK 540
QY 541 DMVRTGLLNLNGVLLLSLAMNTWAQAI FOLGTFPDWANTHAANTVATLPPALTNNVTQTL 600
Db 541 DMVRTGLLNLNGVLLLSLAMNTWAQAI FOLGTFPDWANTHAANTVATLPPALTNNVTQTL 600

RESULT 3
US-10-167-994-4
; Sequence 4, Application US/10167994
; Publication No. US20030082647A1
; GENERAL INFORMATION:
; APPLICANT: Reenan, Robert A.
; APPLICANT: Rogina, Blanka
; APPLICANT: Helfand, Stephen L.
; TITLE OF INVENTION: TRANSPORTER PROTEIN
; FILE REFERENCE: 13407-013001
; CURRENT APPLICATION NUMBER: US/10/167,994
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/255,013
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 10/017,479
; PRIOR FILING DATE: 2001-12-12
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 602
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-167-994-4

Query Match 87.0%; Score 2691; DB 14; Length 602;
Best Local Similarity 85.9%; Pred. No. 7.6e-237;
Matches 517; Conservative 41; Mismatches 42; Indels 2; Gaps 2;
QY 1 MAALAAKAKWSARLLVLLVPLALLPILFALPKGRCGLYVILLMAVWCTEALPLS 60
Db 1 MAALAAKAKWSARLLVLLVPLALLPILFALPKGRCGLYVILLMAVWCTEALPLS 60
QY 61 VTALLPILFPPMGILPSSKVCPOYFLDTNFIPLSGLMASAIEERNHRRIALKVLMLV 120
Db 61 VTALLPILFPPMGILPSSKVCPOYFLDTNFIPLSGLMASAIEERNHRRIALKVLMLV 120
QY 121 GVQPARLILGMVTTSFSLMNLSTASTAMMLPIASAILKSLFGQDTRKOLPREGDST 180
Db 121 GVQPARLILGMVTTSFSLMNLSTASTAMMLPIANAILKSLFGQKEVRKOPSQSEENT 180
QY 181 AAVRNGRLTVPTENQFLASSEG-CHAEDEVAPLELPDPS-KEEHRNINWKGFLISIPY 238
Db 181 AAVRNGRLTVPTENQFLASSEG-CHAEDEVAPLELPDPS-KEEHRNINWKGFLISIPY 240
QY 239 SASIGGTATLTGTAPNLILGQKSFPPQCDVNVFGSWFIFAPPLMLLFLVGLWLSIFL 298
Db 241 SASIGGTATLTGTAPNLILGQKSFPPQCDVNVFGSWFIFAPPLMLLFLVGLWLSIFL 300
QY 299 YGMSWRGRKNSKLDQVAEDKAKAVIOEBFQNLGPDKFAEQAVFIFLCLFALLFSRD 358
Db 301 YGGLSFRGRKNSKSEIRTNADRAVARIREYQNLGPDKFAEQAVFIFLCLFALLFSRD 360
QY 359 PKFIFGNASLFPAGFVSDAVTGVAVITILFFPPSQPSLKWDFDKAPNSETTEPLLSWK 418
Db 361 PKFIFGNASLFPAGFVSDAVTGVAVITILFFPPSQPSLKWDFDKAPNSETTEPLLSWK 420
QY 419 AQETVPWNIILLGGGFAMAKCESGLSAGWIGQLHPLHVPPLAVLLITVVIAFTTE 478
Db 421 AQETVPWNIILLGGGFAMAKCESGLSAGWIGQLHPLHVPPLAVLLITVVIAFTTE 480
QY 479 FASNTATIIIFLPVLAELAIRLHVHPLYLMPGTVCSCYAFMLPVSTPPNSIAFSTGHLL 538

Db 481 FASNTATIIIFLPVLAELAIRLHVHPLYLMPGTVCSCYAFMLPVSTPPNSIAFASGHLL 540
QY 539 VKDMVRTGLLNLNGVLLLSLAMNTWAQAI FOLGTFPDWANTHAANTVATLPPALTNNVTQ 598
Db 541 VKDMVRTGLLNLNGVLLLSLAMNTWAQAI FOLGTFPDWANTHAANTVATLPPALTNNVTQ 600
QY 599 TL 600
Db 601 TL 602

RESULT 4
US-10-108-260A-3862
; Sequence 3862, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1el full length cdna
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3862
; LENGTH: 520
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-3862

Query Match 75.2%; Score 2326.5; DB 15; Length 520;
Best Local Similarity 81.4%; Pred. No. 1.3e-203;
Matches 452; Conservative 36; Mismatches 30; Indels 37; Gaps 3;
QY 48 MAVWCTEALPLSVTALLPILFPPMGILPSSKVCPOYFLDTNFIPLSGLMASAIEERN 107
Db 1 MAVWCTEALPLSVTALLPILFPPMGILPSSKVCPOYFLDTNFIPLSGLMASAIEERN 60
QY 108 LHRRIALKVLMLVGVQPARLILGMVTTSFSLMNLSTASTAMMLPIASAILKSLFGQD 167
Db 61 LHRRIALKVLMLVGVQPARLILGMVTTSFSLMNLSTASTAMMLPIANAILKSLFGQKE 120
QY 168 TRKOLPREGDSTAAVRNGRLTVPTENQFLASSEG-CHAEDEVAPLELPDPS-KEEHR 225
Db 121 VRKOPSQSEENTAAVRNGRLTVPTENQFLASSEG-CHAEDEVAPLELPDPS-KEEHR 180
QY 226 RNIWKGFLISIPYASIGGTATLTGTAPNLILGQKSFPPQCDVNVFGSWFIFAPPLML 285
Db 181 RNIWKGFLISIPYASIGGTATLTGTAPNLILGQKSFPPQCDVNVFGSWFIFAPPLML 240
QY 286 LFLVGLWLSIFLGGMSWRGRKNSKLDQVAEDKAKAVIOEBFQNLGPDKFAEQAVFI 345
Db 241 LFLVGLWLSIFLGGMSWRGRKNSKLDQVAEDKAKAVIOEBFQNLGPDKFAEQAVFI 292
QY 346 LFLVGLWLSIFLGGMSWRGRKNSKLDQVAEDKAKAVIOEBFQNLGPDKFAEQAVFI 405
Db 293 LFLVGLWLSIFLGGMSWRGRKNSKLDQVAEDKAKAVIOEBFQNLGPDKFAEQAVFI 325
QY 406 PNSTEPLLSWKKAQETVPWNIILLGGGFAMAKCESGLSAGWIGQLHPLHVPPLLA 465
Db 326 PNSTEPLLSWKKAQETVPWNIILLGGGFAMAKCESGLSAGWIGQLHPLHVPPLLA 385
QY 466 VLLITVVIAFTTEPASNTATIIIFLPVLAELAIRLHVHPLYLMPGTVCSCYAFMLPVST 525
Db 386 VLLITVVIAFTTEPASNTATIIIFLPVLAELAIRLHVHPLYLMPGTVCSCYAFMLPVST 445
QY 526 PPSNIAFSTGHLLVKDMVRTGLLNLNGVLLLSLAMNTWAQAI FOLGTFPDWANTHAANTV 585
Db 446 PPSNIAFSTGHLLVKDMVRTGLLNLNGVLLLSLAMNTWAQAI FOLGTFPDWANTHAANTV 505
QY 586 TALPPALTNNVTQTL 600
Db 506 TALPPALTNNVTQTL 520

APPLICANT: Spaderna, Steven K.
APPLICANT: Catterton, Elina
APPLICANT: Leite, Mario W.
APPLICANT: Zhong, Haihong
APPLICANT: Alsobrook, John P.
APPLICANT: Lepley, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Burgess, Catherine E.
TITLE OF INVENTION: No. US20040043302A1el Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-290C
CURRENT APPLICATION NUMBER: US/10/092,900A
CURRENT FILING DATE: 2002-03-07
PRIOR APPLICATION NUMBER: USSN 60/274,322
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: USSN 60/283,675
PRIOR FILING DATE: 2001-04-13
PRIOR APPLICATION NUMBER: USSN 60/338,092
PRIOR FILING DATE: 2001-12-03
PRIOR APPLICATION NUMBER: USSN 60/274,281
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: USSN 60/274,191
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: USSN 60/325,681
PRIOR FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: USSN 60/304,354
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: USSN 60/279,995
PRIOR FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: USSN 60/294,899
PRIOR FILING DATE: 2001-05-31
PRIOR APPLICATION NUMBER: USSN 60/287,424
PRIOR FILING DATE: 2001-04-30
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 768
SEQ ID NO 48
LENGTH: 616
TYPE: PRT
ORGANISM: Homo sapiens
US-10-092-900A-48

Query Match 45.7%; Score 1412.5; DB 15; Length 616;
Best Local Similarity 46.6%; Pred. No. 6.4e-120;
Matches 292; Conservative 102; Mismatches 171; Indels 61; Gaps 13;

Qy 4 LAALAKKWSARRLLVLLVPLALLPILFALPKKGRCLYVILLMAVYVCTEALPLSVTA 63
Db 1 MASALSYVSKPSVILFVTLPLLLPILVILPAKVSVCAYVILLMAVYVCTEVIPLAVTS 60

Qy 64 LLPILFPFMGILPSSKVCQYFLDTNPLFLSGLIMASAEERNLHRRALKVLMVGVQ 123
Db 61 LMPVLLPFLQILDSRQVCQYMKDTNMLFLGLLIVAVAVERNLHRRALKVLMVGVQ 120

Qy 124 PARLILGMVTTSLSMWLSNTASTAMMLPIASAILK-----SLFGQDTRKDL 172
Db 121 PARMLGFMGTALLSMWISNTATTAMVPIVEAILQOMEATSAAEAGLEGQGTINNL 180

Qy 173 PREGEDSTAAVRGNGLRTVPTMQLFASSEGHAEDVEAPLE-----LPDD 218

Db 412 TLLFVPSQKPK-----FNFRSQTEGSKSPVLIAPPPLLDWKVTOBKVPWGVLLGGGFA 467
Qy 437 MAKGCEBSGLSAGIQLHPLHLEHPVPLLAVALITVIAFFTEPASNTATIIIFLPLVLAEL 496
Db 468 LAKGEASGLSVWKGQKWEPLHAPVPAITLILSLVAVFECTSNVATITLFLFIFASM 527
Qy 497 AIRLHVHPLYLMIPGTVSCSVAFMLPVSTPPNSIAFSTGHLLVKDMVTRTGLLMNLMGVLL 556
Db 528 SRSLGNPLYMLPCTLLSASFAFMLPVATPBNVAVFTYGLHKVADMVKTGVTIMNIIGVFC 587
Qy 557 LSLAMNTWAQAIFFOLGTTPDWN--TH 581
Db 588 VFLAVNTWGRAIFDLDFHFPDWNVTH 613

RESULT 7
US-10-403-161-68
; Sequence 68, Application US/10403161
; Publication No. US20040043930A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David et al.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-573C
; CURRENT APPLICATION NUMBER: US/10/403,161
; PRIOR FILING DATE: 2003-03-31
; PRIOR APPLICATION NUMBER: 60/370349
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/384543
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/370969
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 60/403748
; PRIOR FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: 60/372019
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/374379
; PRIOR FILING DATE: 2002-04-22
; PRIOR APPLICATION NUMBER: 09/779679
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/181045
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: 10/055877
; PRIOR FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: 60/262892
; PRIOR FILING DATE: 2001-01-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: Curaseq1ist version 0.1
; SEQ ID NO 68
; LENGTH: 616
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-403-161-68

Query Match 45.7%; Score 1412.5; DB 15; Length 616;
Best Local Similarity 46.6%; Pred. No. 6.4e-120;
Matches 292; Conservative 102; Mismatches 171; Indels 61; Gaps 13;

Qy 4 LAALAKKWSARRLLVLLVPLALLPILFALPKKGRCLYVILLMAVYVCTEALPLSVTA 63
Db 1 MASALSYVSKPSVILFVTLPLLLPILVILPAKVSVCAYVILLMAVYVCTEVIPLAVTS 60

Qy 64 LLPILFPFMGILPSSKVCQYFLDTNPLFLSGLIMASAEERNLHRRALKVLMVGVQ 123
Db 61 LMPVLLPFLQILDSRQVCQYMKDTNMLFLGLLIVAVAVERNLHRRALKVLMVGVQ 120

Qy 124 PARLILGMVTTSLSMWLSNTASTAMMLPIASAILK-----SLFGQDTRKDL 172
Db 121 PARMLGFMGTALLSMWISNTATTAMVPIVEAILQOMEATSAAEAGLEGQGTINNL 180

Qy 173 PREGEDSTAAVRGNGLRTVPTMQLFASSEGHAEDVEAPLE-----LPDD 218

Db 181 NAEEDDTKAVLGG--KCVAIISYVKKVEKLIQINNLMTPKLEKQEQDGLGIRPD 238
QY 219 S---KESEHRNINWGLFISIPYSASIGGTATLTGTAPNLLILGOLKSPFQOC-DVNVFG 274
Db 239 SAQCOEDQERKRLCKAWTLCICYAASIGGTATLTGTGNVLLGOMNELPDSKDLVNF 298
QY 275 SWFIFAPPLMLLFLVGLWMLFISFYGMSW-RGW-----RKNKSKLDQVAEDKAKAVIQE 328
Db 299 SWFAFAFPMVLMVLLFAWMLQFYMFSSFKSGCGLESKQNEK-----AALKVLQ 351
QY 329 EFQNLGPIKFAQAVFILPCLFAILLRSRDPKFTPGWASL-FAPG-----FVSDAVTGVAVI 384
Db 352 EYRKLGLSFAEINVLICFFLLVILWFSDPGFMPGLTVAWVEGETKYSVDATVAIFVA 411
QY 385 TILFFFSQKPSLKWDFDKAPNSETE-----PLLSMKKAQETVPWNIIILGSGGPA 436
Db 412 TLLFIVSQPK-----FNFSQTEEGSPVLIAPPPLLDWKVTQKVPWGIVLLGGGPA 467
QY 437 MAKCEBSGLSAMTGGQLHPLHVPPLVLLVIAFFTEFASNTATIIIFLPVLAEL 496
Db 468 LAKGSEASGLSVWVKQWQEPHAPVPAATILSLVAVFTECTSNVATTTFLPIFASM 527
QY 497 AIRLHVPLIYMIPTGVSVCYAFMLPVSTPNSIAPSTGHLVVKDMVRTGLMLNMGVLL 556
Db 528 SRSIGLNPXYIMLPCTLSASFAFMLPVATPPNAIVFTYGHKLVADMVKTGVIMNIIGVFC 587
QY 557 LSLAMNTWAQAIQOLGTFPDWAN-TH 581
Db 588 VFLAVNTWGRAIFDLHDHPPDANVTH 613

RESULT 8
US-10-173-519-2
; Sequence 2, Application US/10173519
; Publication No. US20020193582A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
; TITLE OF INVENTION: 69624, A Human Transporter Family Member
; TITLE OF INVENTION: and Uses Therefor
; FILE REFERENCE: MF01-098PIRNM
; CURRENT APPLICATION NUMBER: US/10/173,519
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 60/298,970
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 568
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-173-519-2
Query Match 45.6%; Score 1411.5; DB 13; Length 568;
Best Local Similarity 47.1%; Pred. No. 7.1e-120;
Matches 286; Conservative 100; Mismatches 150; Indels 71; Gaps 13;

QY 4 LAALAKVWSARRLLVLLVPLALPILFALPKGECGLYVILLMAVYVWCTEALPLSVTA 63
Db 1 MASALSVSFKPSVILFVTPLLPLVILMPAKFVRCVAVIILMAIYVCTEVIPLAVTS 60
QY 64 LLPILFPFGIIPSSKVCQYFIDTNTFLSGLIMASAEERNLHRRIRKALVKVLMVGVQ 123
Db 61 LMPVLLPFLQILDQRVCYQVKOTNNLFLGLIIVAVAVERNWNLHRIALRTLLWVGAK 120
QY 124 PARILGMVTTSPLSMWLNTASTAMMLPIASAILKSLFGQDRDKDLPREGEDSTAAV 183
Db 121 PARMLGMFGVMTALLSMWISNTATTAMVPIVEAILQOM-----EATSA 165
QY 184 RGNGLRVTPTMQPLASSEGHAEDVEAPLELPPD-----SKEEHRNRINWGLF 233
Db 166 TEAGLELV-----DKGAK-----ELPGSQVIFGPTLGGQEDQERKRLCKAWT 209

QY 234 ISIPYSASIGGTATLTGTAPNLLILGOLKSPFQOC-DVNVFGSWFIFAPPLMLLFLVGV 292
Db 210 LCICVAASIGGTATLTGTGNVLLGOMNELPDSKDLVNFASWFAFPMVLMVLLFAW 269
QY 293 LMSIFLYGMSW-RGW-----RKNKSKLDQVAEDKAKAVIQEFGNLGPIKFAQAVFIL 346
Db 270 LWLQFYMYRNFKSKGCGLESKQNEK-----AALKVLQBEYRKLGLSFAEINVLIC 322
QY 347 FCLFAILLRSRDPKFTPGWASL-FAPG-----FVSDAVTGVAVIITLFFFSQKPSLKWDF 402
Db 323 FELLVILWFSDPGFMPGLTVAWVEGETKYVDATVAIFVATLTLFIVSQPK-----FN 378
QY 403 FKAPNSETE-----PLLSMKKAQETVPWNIIILGSGGFAMAKGCEBSGLSAWIGQQLH 455
Db 379 FRSQTEERKTPYPPPLLDWKVTQKVPWGIVLLGGGFALAKGSEASGLSVWVKQME 438
QY 456 PLEHVPPLVAVLITVIAFFTEFASNTATIIIFLPVLAELAIRLHVPLIYMIPTGVS 515
Db 439 PLHAVPPAAITILSLVAVFTECTSNVATTTFLPIFASMSRSIGLNPYIMLPCTLSA 498
QY 516 SYAFMLPVSTPNSIAPSTGHLVVKDMVRTGLMLNMGVLLSLAMNTWAQAIQOLGTP 575
Db 499 SFAFMLPVATPPNAIVFTYGHKLVADMVKTGVIMNIIGVFCVFLAVNTWGRAIFDLHDH 558
QY 576 DWAN-TH 581
Db 559 DWANVTH 565

RESULT 9
US-10-403-161-72
; Sequence 72, Application US/10403161
; Publication No. US20040043930A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David et al.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-573C
; CURRENT APPLICATION NUMBER: US/10/403,161
; CURRENT FILING DATE: 2003-03-31
; PRIOR APPLICATION NUMBER: 60/370349
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/384543
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/370969
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 60/403748
; PRIOR FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: 60/372019
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/374379
; PRIOR FILING DATE: 2002-04-22
; PRIOR APPLICATION NUMBER: 09/779679
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/181045
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 10/055877
; PRIOR FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: 60/262892
; PRIOR FILING DATE: 2001-01-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 72
; LENGTH: 568
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-403-161-72

Query Match 45.6%; Score 1411.5; DB 15; Length 568;
Best Local Similarity 47.1%; Pred. No. 7.1e-120;
Matches 286; Conservative 100; Mismatches 150; Indels 71; Gaps 13;
QY 4 LAALAKVWSARRLLVLLVPLALPILFALPKGECGLYVILLMAVYVWCTEALPLSVTA 63

```

Db      1 MASALSYSVSKFSFVILFVTPLLLPLVILMPAKFVRCAVYVILMAIYVCTEVIPLAVTS 60
Qy      64 LPLIILFFPMGILPSSKVCPOYFLDTNPLFLSGLIMASAIERNLHRRALKVLMLGVQ 123
Db      61 LMPVLLPFLFQILDSRQVCQYKMDTNMLFLGGLIVAVAVERNLHRRALKVLMLGVQ 120
Qy     124 PARLILGMVVTTSPLSMWLSNTASTAMMLPIASAILKSLFGQDRTRKOLPREGEDSTA 183
Db     121 PARMLGFMGTALLSMWISNTATTAMVPIVEAILQOM-----EATSAA 165
Qy     184 RGNGLRTVPTMQFLASSEGHAEDVEAPLEPDD-----SKEEHRRNTWKGL 233
Db     166 TEAGLELV-----DKGKAK-----ELPGSQVIFEGTTLGQOQDERKRLCKAMT 209
Qy     234 ISIPYSASIGGTATLTGTAPNLILGQLKSFPPQC-DVNFSGSWFIFAPPLMLLPLLVGW 292
Db     210 LCICYAASIGGTATLTGTGNVLLGQNNELFPDSKDLVNFASFAPFAPNMLVLLFAW 269
Qy     293 LWSIFLYGMSW-RGW-----RKNSKLDQVAEDKAKAVIOEFQNLGPIKPAEQAVFIL 346
Db     270 LWLQVYMRFNFKKSWGCGLESKNEK-----AALKVLQBEYRKLGLPLSFAEINVLIC 322
Qy     347 FCLFALLFSRDPKFIQWASL-FAPG---FVSDAVTGAIVTILFFPSPKPSLKWFD 402
Db     323 FFLVILWFSRDPGMPGWLTVAVVEGETKYVSDATVAIFVATLLFIVPSQPKP----FN 378
Qy     403 FKAPNSET-----PLLSWKKAQETVPMNIIILLGGGFAMAKGCEBSGLSANTGGQLH 455
Db     379 FRSQTEERKTPFPFPPDLLDKWVQKVPWGIIVLLGGFALAKGSEASGLSVNMGKQME 438
Qy     456 PLEHVPPLAVLLITVIAFTPEFASNTATIIIFLPVLAELAIRLHVHPLYLMIPGTVSC 515
Db     439 PLHAVPPAAITLILSLVAVFECTSNVATTLFLPIFASMSRSIGLNPILMPLCTLSA 498
Qy     516 SYAFMLPVSTPNSTAFSTGHLVKNMVRTGLLNMGLVLLSLAMNTWAQAIPLGTFPP 575
Db     499 SFAFMLPVATPPNAIVFTYGHVKVADMVKTGVMNIIIGVFCVFLAVNTWGRAIFDLDHFP 558
Qy     576 DWAN-TH 581
Db     559 DWANVTH 565

```

RESULT 11

```

US-10-718-359-6
; Sequence 6, Application US/10718359
; Publication No. US20050095240A1
; GENERAL INFORMATION:
; APPLICANT: MEDICAL COLLEGE OF GEORGIA RESEARCH INSTITUTE, INC.
; TITLE OF INVENTION: NaCT AS A TARGET FOR LIFESPAN EXPANSION AND WEIGHT REDUCTION
; FILE REFERENCE: 275.0008 0101
; CURRENT APPLICATION NUMBER: US/10/718,359
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/428,469
; PRIOR FILING DATE: 2002-11-22
; PRIOR APPLICATION NUMBER: 60/459,441
; PRIOR FILING DATE: 2003-04-01
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 568
; TYPE: PRT
; ORGANISM: human NaCT
US-10-718-359-6

```

Query Match 45.6%; Score 1411.5; DB 17; Length 568;
 Best Local Similarity 47.1%; Pred. No. 7.1e-120;
 Matches 286; Conservative 100; Mismatches 150; Indels 71; Gaps 13;

```

Qy      4 LAALAKKWSARRLLVLLVPLALLPILFALPKPEGRCCLYVILLMAVYVCTEALPLSVTA 63
Db      1 MASALSYSVSKFSFVILFVTPLLLPLVILMPAKFVRCAVYVILMAIYVCTEVIPLAVTS 60

```

RESULT 10

```

US-10-490-080-1
; Sequence 1, Application US/10490080
; Publication No. US20040253597A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.
; TITLE OF INVENTION: Novel Protein and its DNA
; FILE REFERENCE: P02-0109PCT
; CURRENT APPLICATION NUMBER: US/10/490,080
; CURRENT FILING DATE: 2004-03-17
; PRIOR APPLICATION NUMBER: JP 2001-281992
; PRIOR FILING DATE: 2001-09-17
; PRIOR APPLICATION NUMBER: JP 2001-306873
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: JP 2002-113279
; PRIOR FILING DATE: 2002-04-16
; NUMBER OF SEQ ID NOS: 42
; SEQ ID NO 1
; LENGTH: 568
; TYPE: PRT
; ORGANISM: Human
US-10-490-080-1

```

Query Match 45.6%; Score 1411.5; DB 16; Length 568;
 Best Local Similarity 47.1%; Pred. No. 7.1e-120;
 Matches 286; Conservative 100; Mismatches 150; Indels 71; Gaps 13;

```

Qy      4 LAALAKKWSARRLLVLLVPLALLPILFALPKPEGRCCLYVILLMAVYVCTEALPLSVTA 63

```


TYPE: PRT
ORGANISM: Xenopus laevis
US-09-729-094-4

Query Match 45.4%; Score 1405.5; DB 9; Length 619;
Best Local Similarity 45.1%; Pred. No. 2.8e-119;
Matches 280; Conservative 122; Mismatches 170; Indels 49; Gaps 11;

Query Match 45.4%; Score 1405.5; DB 14; Length 619;
Best Local Similarity 45.1%; Pred. No. 2.8e-119;
Matches 280; Conservative 122; Mismatches 170; Indels 49; Gaps 11;

QY	4	LAALAKKWSARRLLVLLVPLALPILFALPPKGRCLYVLLMAVYVCTEALPLSVTA	63
DB	1	MVSIQKWLARNYFIIFLPLPLVPTKEASCGFVIIVMALFWCTEALPLAVTA	60
QY	64	LLPILFFPMGILPSSKVCPOVFLDTNLFSLGLMASAIEERNLHRRALVKVLMVGVO	123
DB	61	LFPVLLFFPMGIMDSTAVCSQYLKDTNMLFICGLLVAISVEKNLHRRALVKVLMVGVO	120
QY	124	PARLILGMVMTTSLFSLMWSNTASTAMMLPIASAILKSLF-----GORDTRKDL	172
DB	121	PALLILGFVAVTAFSLMWSNTASTAMMLPIASAILKSLF-----GORDTRKDL	170
QY	173	PREGEDSTAARVGNGLRTVPTMQFLASSEGHAEDVE-----APLEL-PD	217
DB	181	NGMENDMTESVMPSGMALAIDNTVATENEGFEIOEKSTKDPSPKQSGIGPVIPE	240
QY	218	DSKEEHR-----NIWKGFLISIPYSASIGTATLTGTAPNLILLGQKSPFPQ-CDVV	271
DB	241	DEKQTEEKQKHLKICKGMSLCVCYSASIGGIATLTGTPNLVKGQMDLFPENNII	300
QY	272	NFGSWFIAPPLMLFLVGLWISFLYGMWSRGRKNSKLQDVAEDKAK-----AVI	326
DB	301	NFASWFGFAPPTMLVLLSMLWLFQIYLVNF-----KNFGCGGNASQKQKRAFRVI	356
QY	327	QEEFNQGLPIFAQVAFILFCLFAILLFSRDPKFIQWASL-FAPG---FVSDAVTGA	382
DB	357	SGEHKLGSMTFAISVLVFLVLLVFTREPFGMPGWATISFNKGGKEMVTDATVAF	416
QY	383	IVTILFFPPSPKSLKWFDFK---APNSETP-LLSWKKAQETVPMNIIILLGGGFAMA	438
DB	417	VSLMFFFPSPKSLKWFDFK---APNSETP-LLSWKKAQETVPMNIIILLGGGFAMA	476
QY	439	KGCEESGLSAMIGGQLHLEHVPPLVLLVITVIAFTEFASNTATIIIFLPVLAELAI	498
DB	477	KGCEESGLSMLGKLTPLQSIIPPAALILCLLVATFECTSNVATTTFLPILASMAK	536
QY	499	RLVHPLYLIMPTGVSCSYAFMLPVSTPNSIAFSTGHLLVKDMVRTGLLMLMGLVLLS	558
DB	537	AIQLNPLYLIMPTGVSCSYAFMLPVSTPNSIAFSTGHLLVKDMVRTGLLMLMGLVLLS	558
QY	559	LAMNTWAQAIQFOLGTFPDWAN	579
DB	597	LAINSWGFMFNLGTFPDSWAN	617

RESULT 15

US-10-092-900A-44
Sequence 44, Application US/10092900A
Publication No. US20040043382A1
GENERAL INFORMATION:
APPLICANT: Padigaru, Muralidhara
APPLICANT: Spvtek, Kimberly A.
APPLICANT: Shenoy, Sureeh G.
APPLICANT: Taupier Jr., Raymond J.
APPLICANT: Pena, Carol E.A.
APPLICANT: Li, Li
APPLICANT: Zehusen, Bryan D.
APPLICANT: Gusev, Vladimir Y.
APPLICANT: Ji, Weizhen
APPLICANT: Gorman, Linda
APPLICANT: Miller, Charles E.
APPLICANT: Kekuda, Ramesh
APPLICANT: Patturajan, Meera
APPLICANT: Gangolli, Esha A.
APPLICANT: Vernet, Corine A.M.
APPLICANT: Guo, Xiaojia Sasha T.
APPLICANT: Tchernev., Velizar T.
APPLICANT: Fernandes, Elma R.
APPLICANT: Casman, Stacie J.

RESULT 14

US-10-435-631-4
Sequence 4, Application US/10435631
Publication No. US20030186381A1
GENERAL INFORMATION:
APPLICANT: CHATURVEDI, Kabir et al
TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
TITLE OF INVENTION: AND USES THEREOF
FILE REFERENCE: CLO00662CON
CURRENT APPLICATION NUMBER: US/10/435,631
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 619
TYPE: PRT
ORGANISM: Xenopus laevis
US-10-435-631-4

APPLICANT: Malvankar, Uriel M.
APPLICANT: Gerlach, Valerie
APPLICANT: Liu, Yi
APPLICANT: Anderson, David W.
APPLICANT: Spaderna, Steven K.
APPLICANT: Catterton, Elina
APPLICANT: Leite, Mario W.
APPLICANT: Zhong, Haihong
APPLICANT: Alsobrook, John P.
APPLICANT: Lepley, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Burgess, Catherine E.
TITLE OF INVENTION: No. US20040043382A1e1 Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-290C
CURRENT APPLICATION NUMBER: US/10/092,900A
CURRENT FILING DATE: 2002-03-07
PRIOR APPLICATION NUMBER: USSN 60/274,322
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: USSN 60/283,675
PRIOR FILING DATE: 2001-04-13
PRIOR APPLICATION NUMBER: USSN 60/338,092
PRIOR FILING DATE: 2001-12-03
PRIOR APPLICATION NUMBER: USSN 60/274,281
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: USSN 60/274,191
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: USSN 60/325,681
PRIOR FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: USSN 60/304,354
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: USSN 60/279,995
PRIOR FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: USSN 60/294,899
PRIOR FILING DATE: 2001-05-31
PRIOR APPLICATION NUMBER: USSN 60/287,424
PRIOR FILING DATE: 2001-04-30
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 768
SEQ ID NO 44
LENGTH: 568
TYPE: PRT
ORGANISM: Homo sapiens
US-10-092-900A-44

Query Match 45.3%; Score 1400.5; DB 15; Length 568;
Best Local Similarity 47.0%; Pred. No. 7.2e-119; Indels 73; Gaps 14;
Matches 286; Conservative 100; Mismatches 149;
QY 4 LAALAKKVVSGARRLLVLLVPLALLPILFALPPKRGCLYVILLMAVYWCTEALPLSVTA 63
DB 1 MASALSIVSKFSFVILFVTPULLPLVILMPAKVS-CAYVILMAIYWCTEVPLAVTS 59
QY 64 LLPILPPFMGILPSSKVCQYFLDTNFIPLSLGLIMASAIERNLHRRIALKVLMLGVQ 123
DB 60 LMPVLLPFLQILDQRQVCQYKMDTNMLFLGGLIVAVAVERNLHRRIALRLLMWGAK 119
QY 124 PARLLGMMVTTSLMNLSTANMLPLASAILKSLFGQDRDKLPREGEDSTAAY 183
DB 120 PARMLGMFGVTALLSWMISNTATAMVPVIVEAILQOM-----EATSA 164
QY 184 RGNGLRVTPTMQPLASSEGHAEDVEAPLELPD-----SKEEHRNRNIWKGF 233
DB 165 TEAGLELV-----DKGKAK-----ELPGSQVIFSGPTLGGQDQERKLCCKAMT 208
QY 234 ISIPYSASIGGTATLTGTAPNLILLGQKSFPOC-DVNFSGSWFIFAFPLMLLFLVGV 292
DB 209 LCICYAASIGGTATLTGTGNVLLGQNELFPDSDLVNFASFAPAFPNMLVMLLFAW 268
QY 293 LWISFLYGGMSW-RGW-----RKNSKLQDVAEDKAKAVIQEERFONLGPICKFAEQAVF 346
DB 269 LWLFVYVNFSSFKKSWCGGLESKNEK-----AALKVIOEEYRKLGPISFAEINVLC 321
QY 347 FCLFAILLFSRDPKFIQWASL-FAPG---FVSDAVTGAIVTILFPFPQKPSLKWWD 402

Db 322 FELLVILWFSRDPGFMGMLTVAWVEGETKYVSDATVAIFVATLLFIVPSQPK-----FN 377
QY 403 PKAPNSET-----PLLSWKKAQETVPKNIILLGGGFAMAKGCBESGLSAMIGQOL 454
Db 378 FRSQTEGKSPVLIAPPPLLDWKVTQEKVPMGIVLLGGGFALAKGSEASGLSVWVGKQM 437
QY 455 HPLEHVPPLLAVALLITVVIATFEPASNTATIIIFLPVLAELAIRLHVHPLYLMIPTVS 514
Db 438 EPLHAVPPAAITLILSLVAVFECTSNVATITLFLPIFASMSRISGLNPLYIMLPCTLS 497
QY 515 CSYAFMLPVSTPPNSIAFSTGHLLVKDMVTRGLLMLNMGVLLLSLAMNTWAAIQTGTF 574
Db 498 ASFAFMLPVATPPNAIVFTYGLKQVADMVKTGVIIMNIIGVFCVFLAVNTWGRAIFDLDF 557
QY 575 PDWAN-TH 581
Db 556 PDWANVTH 565

Search completed: June 30, 2005, 09:32:52
Job time : 79.8579 secs